EPA/600/R-94/204a
October 1992

HAZARD IDENTIFICATION IN CARCINOGEN RISK ANALYSIS: AN INTEGRATIVE APPROACH

by

Douglas J. Crawford-Brown
Kenneth G. Brown

October, 1992

This project report is part of the U. S. EPA's program of *Research to Improve Health Risk Assessment*. It was funded by the Office of Health and Environmental Assessment, Office of Research and Development, U. S. Environmental Protection Agency, Washington, D. C. 20460.

DISCLAIMER

Although the information in this document has been funded by the U.S. Environmental Protection Agency, under contract 68-C9-0009, Work Assignment S-1-56 (Kenneth G. Brown, Ph.D., Inc.), it does not necessarily reflect the views of the Agency and no official endorsement should be inferred. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

Contents (continued)

TABLES

WORKING TABLES

FIGURES

PREFACE

This report was supported by the U.S. Environmental Protection Agency, Office of Health and Environmental Exposure Assessment, as part of the Agency's program in "Research to Improve Health Risk Assessment". Part I of the report describes the philosophical foundations for rationality and reasoning available to the analyst for risk assessment. Part II develops a framework to aid the analyst in systematically assembling and evaluating all observational information potentially relevant to forming a judgment of carcinogenicity for an agent of interest.

The strength and type of evidential support for warranting claims of carcinogenicity from observational evidence in Part II derives from the foundational principles discussed in Part I. By implementing these principles in a step-by-step procedure the analyst is systematically guided toward developing a weight-of-evidence judgment of carcinogenicity. *Human judgment* is a key element required throughout. The stepwise procedure is an aid to (not a substitute for) forming judgments based on the available evidence. Individuals may make different judgments from the same observational evidence, particularly if they differ in their "intellectual obligation" regarding types of relevance strategies deemed necessary for warranting claims of carcinogenicity. Differences of opinions, knowledge, and perspectives that individuals bring to bear on evaluating evidence and making judgments will also contribute to diversity. Rational discourse between persons who have formed their individual weight-of-evidence judgments on claims of carcinogenicity is recommended to form a "belief" (e.g., an Agency classification regarding carcinogenicity). If persons have followed the same procedural steps in arriving at different judgments of carcinogenicity, then the sources of their differences are readily identifiable for discussion directed toward conflict resolution.

The task of hazard identification is unique to each agent that might be considered. The framework suggested is intended to contain essential features common to hazard identification of any agent, but also to be sufficiently flexible to accommodate the wide diversity of situations that may arise without excluding useful information. It is anticipated, however, that experience with actual applications will indicate areas in which improvements can be made. In this sense the procedure suggested should be viewed as a prototype. Perhaps the critical test will pertain to the structural capability of the framework to accommodate changes and revisions that may be helpful, particularly to correct any oversights in its development.

EXECUTIVE SUMMARY

This report is directed at improving health risk assessment, and deals exclusively with the hazard identification step of risk analysis. It does not, however, begin with the guidelines of EPA or another organization and attempt to make refinements that may result in reduced uncertainty. Instead, it begins at a more fundamental level, first considering principles of rationality by which to formulate judgments about carcinogenicity from the available information base (the topic of Part I), then addressing what types and sources of information should be assimilated into the informational base and how they may be systematically assimilated and integrated into the formation of a judgment about carcinogenicity (the subject of Part II). The term "carcinogenicity" is expanded from its usual meaning to distinguish within a taxonomy of modes of action, such as initiator, promoter, risk modifier, etc. Additionally, carcinogenicity is not considered to be a property of a chemical; it requires a reference context to be meaningful, e.g., Chemical A is a carcinogen (in some particular sense from among the taxonomy of choices) in Sprague-Dawley rats exposed for life at exposure concentration X. This is an important concept since a chemical may be carcinogenic in one context but not another, and some of the difficult questions to be addressed in hazard identification concern the inter-context extrapolation of evidence, e.g., extrapolation across species or across exposure levels.

To expound further, the two parts of this report are conceptually distinct, yet practically interrelated. Part I contains a review of philosophical ideas that have some bearing on the practical matter of conducting and defending a risk analysis. In particular, discussion is directed toward the issue of rationality with the philosophical literature review focused on a single central question: Under what conditions, and in what sense, can an analyst assert that the claims made in a risk analysis are rational? The significance of *completeness of information* and

coherence of evaluation notwithstanding, the *principle of rationality* is considered fundamental to provide an informed basis for identifying needed sources of information, processing and evaluating it for credibility and relevance, and then making rational judgments. The authors make no claim to having resolved philosophical questions related to the concept of rationality. The review provided in Part I contains, instead, a summary of this topic from the primary literature of philosophy, risk analysis, science, etc.

The discussion in Part I also forms the basis for the development of relevance strategies for assessing evidence in the seven-step framework for performing hazard identification in Part II. Relevance strategies are the bases by which observational information (such as epidemiologic data, or results from chronic animal exposure, *in vitro* tests, or other sources) are warranted as supporting evidence of carcinogenicity. For example, direct empirical observation (I "saw" it) is a relevance strategy that some people may apply to human data demonstrating that cancer incidence increased with exposure to the agent of interest. Empirical correlation (Conditions correlate with situations where the agent was carcinogenic) and theory-based inference (The outcome is consistent with theories of cancer mechanisms) are further examples of relevance strategies. The intrinsic value of various relevance strategies for warranting claims of carcinogenicity may differ between individuals within a rational framework for decision making, thus creating one of the needs for discourse between decision makers.

Human judgment, however, and consequently human discourse to resolve conflicts and work toward an agreement, are considered necessary to hazard identification. To assign a decision rule or other form of computational or analytical scheme to avoid human decision making and potential conflict in resolving differences would impose an artificial and limited framework. It is not possible to foresee all possible contingencies and exigencies in advance,

and even if it were, human decision making and discourse would still be required to formulate the decision rule.

In general, the information available for consideration in forming individual judgments in the hazard identification step may be widely heterogenous and difficult to assess without some structure for guidance. The objective of Part II is to aid this process via a seven-step framework that may be followed by an individual decision maker. An individual assessment of available information guided by a common framework of decision making based on principles of rationality, completeness, and coherence (as discussed in Part I), will lead to a logical conclusion supportable by the individual's own perspective, opinion, and experience, as well as providing a means for comparing the basis of conclusions drawn by different individuals. Assumptions and "judgment calls" are made apparent, indicating areas and degrees of support/uncertainty. This approach should (1) aid in the systematic evaluation of all sources of information that contribute to hazard identification and the overall strength-of-evidence warranted; (2) identify sources of divergence resulting from different perspectives and opinions of individuals, thus contributing to conflict resolution; and (3) help to identify areas for research and assess their potential impact on decision outcome(s) or their level of confidence. The overall objective this project is to develop a logical assemblage of the diverse factors contributing to the decision making process to accomplish these three goals.

FOREWORD

> "In the absence of direct human evidence of carcinogenicity, the conclusion that an agent is liable to cause cancer in man is a matter of judgement...No hard and fast criteria can be laid down that will automatically lead to an appropriate conclusion in all circumstances. Only one rule is absolute: that all the available evidence must always be taken into account." R. Doll (IARC Sci. Pub. No. 65, 1985)

Health risk assessment typically consists of four distinct steps, including (1) Hazard Identification, (2) Dose-Response Assessment, (3) Exposure Assessment, and (4) Risk Characterization. Of interest in this report is the task associated with the first step, hazard identification, wherein the objective is to determine "whether a particular chemical is or is not causally linked with a particular health effect" (NRC, 1983). With respect to cancer, the particular health effect of interest in this report, the EPA guidelines define hazard identification as "a qualitative risk assessment, dealing with the process of determining whether exposure to an agent has the potential to increase the incidence of cancer...The hazard identification component qualitatively answers the question of how likely an agent is to be a human carcinogen" (U.S. EPA, 1986). EPA's approach is to make a judgment based on the weight-of-evidence and to assign the chemical a group classification accordingly. The "evidence" used in the "weight-of-evidence" for hazard identification should include a review of the following information to the extent that it is available: physical-chemical properties and routes and patterns of exposure, structure-activity relationships, metabolic and pharmacokinetic properties, toxicologic effects, short-term tests, and long-term animal studies.

The EPA guidelines also note that there is a need for new methodology not addressed at that time, e.g., the characterization of uncertainty. The present report is part of the agency's subsequent EPA research effort directed at reducing uncertainty in risk analysis, which is currently referred to as research to improve risk assessment. Broadly, there are three sources of

uncertainty intrinsic to risk assessment in general, due to (1) limitations of observational evidence, (2) incomplete knowledge on which to assess the significance of the evidence with respect to predictions of risk, and (3) lack of rigorous development of the process by which current knowledge is applied to available observational evidence to form a conclusion. This report is concerned with the third category of uncertainty, as well as the manner in which the first two categories are reflected in the process of justifying a conclusion of carcinogenicity.

The outcome of the hazard identification step for an agent depends on the evidence available to decision makers and on their means or criteria for assessment of the evidence in forming a judgment regarding carcinogenicity of the agent. It is somewhat analogous to a courtroom trial in which the defendant is the chemical agent charged with causing cancer and the jurors are the decision makers who hear the evidence. The jurors form individual judgments initially and then a collective judgment following discourse. If every juror were expected to weigh the merits of the evidence identically, a single juror would suffice. That, of course, is not the case. Similarly, decision makers may evaluate the same evidence differently. For example, the relative significance of the type of information, such as epidemiologic data, long term animal studies, current theories of carcinogenesis, biological plausibility, physical and chemical properties, and other categories of evidence, may differ between decision makers independent of the strength of the evidence in each category. The axiom that reasonable men may disagree applies to jurors and to decision makers in risk analysis, as elsewhere.

We would replace "reasonable" with "rational" as the objective for decision makers, i.e., the means or criteria for assessment of the evidence in forming a judgment regarding carcinogenicity should be rational. While this objective may appear trivial (After all, who wants decisions that are not rational?), the concept of rationality itself is not at all trivial and needs to be understood if rationality is to be accepted as a guiding principle for judging evidence of

carcinogenicity. This topic is central to Part I of this report entitled *Risk, Regulatory Science, Rationality, and Societal Values*, which contains a review of philosophical ideas and principles likely to have some bearing on the practical matter of conducting and defending a risk analysis. In particular, the material in Part I focuses on the question: Under what conditions, and in what sense, can an analyst assert that the claims made in a risk analysis are rational?

Within this framework of rationality, however, there is a great deal of disagreement as to how the general features should be reflected in specific judgments made by individuals or groups. This disagreement arises from philosophical differences concerning the nature of evidence, how evidence is related to observations and experiments, when evidence is relevant to a specific line of reasoning, when evidence is sufficiently strong to justify a claim, and so on. No attempt is made in this report to resolve these issues, which probably are not resolvable except in the sense of reaching a societal consensus. It is important to bear in mind that selection of a particular view on rationality is as much a matter of human values as of logic, epistemology and procedural rules. These values necessarily enter the discussion since a view on rationality is also a view on human nature and on the form reasoning must take if a belief or claim is to be thought of as having arisen from a process judged worthy of human decisions. In the language of Section 4.1, rationality is both a descriptive and normative goal. The first chapter describes competing conceptions of rationality, and the second chapter formalizes these into judgments and decisions to be made by the analyst(s). The judgments remain an important task of the analyst since the present authors have chosen to avoid imposing their own normative judgments.

If the principles of rationality guide the process of hazard identification, then the informational base available fuels it. As noted in the EPA guidelines and evident in the several categories of evidence listed above, a wide range of evidence is potentially available to the decision maker. The goals of accuracy and minimal uncertainty are both consistent with the

tenet that the best judgment is the most informed judgment. If less than complete information is considered, then uncertainty can be further reduced by expanding the base of information. With respect to accuracy of decisions, incomplete information has the potential to mislead or to bias a decision. Identification of the information useful for hazard identification, a framework to aid in its assessment and integration for decision making, and consideration of assumptions required for extrapolation across contexts of observations (such as between species or dose levels) are addressed in a seven-step guide in Part II of this report, entitled *Assessing Evidential Support for Claims of Carcinogenicity: Essential Elements and a Framework for Application*. This part of the report also addresses the question of what is meant by "carcinogenicity". The concept that an agent is carcinogenic if exposure to it increases the incidence rate of cancer does not cover all possibilities of interest. A taxonomy of carcinogenicity is defined for use instead.

How should the report be read? Here, the answer depends upon the background and interests of the reader. As described above, the report is divided into two parts: the first containing a philosophical discussion concerning rationality and the second consisting of practical matters relating the philosophical principles to the conduct of the hazard identification stage of a risk analysis. Readers already familiar with the literature on rationality, epistemology and philosophy of science might skip Part I entirely, moving directly to Part II for the discussion of applications. Part I contains, however, a number of examples providing insights into how the general philosophical ideas are related to the specific field of risk analysis. Even readers familiar with the literature, therefore, might find it useful to at least scan Part I for the overall flow of ideas and to read the examples provided there.

Those with little background in philosophy are left with two options. It is possible to read Part II on its own since it contains (at least implicitly) consideration of many of the ideas

discussed in Part I. Having mastered the quasi-formal framework of this second chapter, it then would be useful to return to Part I to obtain a more detailed understanding of the various philosophical positions that might be adopted in making the specific judgments called for in the second part. An alternative approach (and probably the best approach) would be to read Chapters 3 and 4 of Part I, which provide a broad discussion of rationality, followed by the entirety of Part II, and finally by the more detailed discussion in Chapter 5 of Part I. In any event, a firm grasp of the material in Part I (and of the primary literature reviewed there) is essential to making full use of the framework in Part II. The need for such a grasp arises from several considerations:

(1) While the quasi-formal framework of Part II contains a distinct set of procedures, these explicit procedures call repeatedly for judgments related to the rationality of collecting, analyzing, judging and employing evidence within lines of reasoning. There are competing conceptions of how these explicit judgments should be made, and these conceptions are reviewed in Part I.

(2) Part I contains material of potential importance in Part II, but not yet formally incorporated into the framework of analysis reported there. The present document is the first product of the authors' broader (and more long-term) research program to link the subject matter of the two chapters. A review of Part I may suggest to the reader new formal considerations to be introduced into the analytic framework of Part II. Similarly, review of Part II may suggest judgments requiring philosophical principles left out of the discussion in Part I. To stimulate the interchange of ideas contained in the separate parts, the authors have included a table (see Table 0) listing the central ideas in Part I and the points at which they become important considerations within the framework introduced in Part II. It is hoped that others will make this table more complete through complimentary research.

(3) In an important sense, the explicit judgments called for in Part II must take place within a context of rational discourse. A central thesis of the entire report is that rationality is more than a formal rule-based procedure for making decisions. It is, instead, a "frame of mind" in which the analyst attempts to determine the quality of philosophical positions concerning beliefs, claims, evidence and evidential reason. To be rational is to carry on a discourse concerning the many potential views on what it means to have well-founded beliefs or claims (this distinction between belief and claim is raised in Part II). Part I contains material essential to such a discourse, even where there is no obvious judgment called for in Part II. The authors have made a strong attempt to show explicitly how the judgments of Part II are related to the principles of rationality set forth in Part I

(See Table 0 mentioned above). Still, there will be points at which the links are not explicit, calling upon the reader to provide creative input into the framework of analysis. This creative input requires a constant and well-reasoned discourse on the philosophical principles drawn from Chapter 1 and , ultimately, from the primary literature.

The overall structure of the report, and of the research program, then is one of:

(a) Developing conceptions of rationality (Chapters 3 and 4 of Part I);

(b) Developing conceptions of evidential reason necessary within all conceptions of rationality (Chapter of Part I);

(c) Formalizing results of (a) and (b) into explicit principles forming an axiomatic base (or set of potential bases) of evidential reasoning (Table 0);

(d) Elucidating the framework of rational discourse concerning claims of hazard identification (Chapters 1 and 5 of Part II);

(e) Developing the bodies of evidence potentially useful in providing the judgments required by hazard identification, and

(f) Making explicit (to the degree possible) the link between (a, b) and (d, e) (Table 0).

PART ONE

RISK, REGULATORY SCIENCE, RATIONALITY AND SOCIETAL VALUES

1. BACKGROUND AND TASK OBJECTIVE

The assessment of the risk of cancer and other adverse health effects associated with exposure to toxic substances is a subject of much debate, posing intricate relations between science and policy. That sentence describes the setting in 1983 when a committee of the National Research Council (NRC) was mandated by a congressional directive to address some fundamental questions regarding the feasibility and merits of alternative modes of conducting risk assessment among all federal regulatory agencies.[59] In particular, the committee was asked to examine whether altered institutional arrangements or procedures can improve regulatory performance. This request was largely motivated by criticisms of risk assessment that ranged broadly from details of the process to administrative management to statutory authority. The committee recommended that regulatory agencies take steps to establish and maintain a clear conceptual distinction between assessment of risk and consideration of risk management alternatives; that is, the scientific findings and policy judgments embodied in risk assessments should be explicitly distinguished from the political, economic, and technical considerations that influence the design and choice of regulatory strategies. The committee also recommended that (1) uniform inference guidelines be developed for the use of federal regulatory agencies in the risk assessment process that would be evaluated regularly for their usefulness and revised as needed, (2) the evolving scientific basis of risk assessment should be critically assessed, and (3) explicit underlying assumptions and policy ramifications of the inference options in each component of the risk assessment process should be made explicit. The NRC committee

differentiated between four components of risk assessment: hazard identification, dose-response assessment, exposure assessment, and risk characterization.

The EPA had already formulated interim guidelines for risk assessment at the time of the NRC committee's report in 1983. The Agency's current risk assessment guidelines adopt the recommended distinction between risk management and risk assessment, and the categorization of the latter into four component elements.[60] The first of these, hazard identification, consists of two elements (long-term animal studies and human studies) that provide the primary data and five elements (physical-chemical properties and routes and patterns of exposure, structure activity relationships, metabolic and pharmacokinetic properties, toxicologic effects, and short-term tests) that may contribute useful information to the qualitative hazard evaluation as available. Aside from at least one central data source from an animal or human study, the information available on the hazard identification categories will vary from chemical-to-chemical depending on the available research base. Consequently, there is no prescribed way of integrating the available information in all cases, either in the evaluation of hazard or the designation of weight-of-evidence classification.

Chemical agents are assessed on a case-by-case basis. Knowledge of metabolic and pharmacokinetic properties, toxicologic effects, and the other categories of hazard identification help to fill informational gaps otherwise bridged (sometimes implicitly) by assumptions and thus reduce uncertainty. Essential to the process of hazard identification and subsequent dose-response assessment when warranted, however, is not only completeness of information but coherence and rational evaluation as well. An individual assessment of available information guided by a common framework of decision making based on these factors will lead to logical conclusion(s) supportable by the individual's own perspective, opinions, and experience, as well as providing a means for comparing the basis of conclusions drawn by different individuals.

Assumptions and "judgment calls" are made apparent, indicating areas and degrees of support/uncertainty (to be replaced by the more definitive term "warrant" in the text). This approach should (1) aid in the systematic evaluation of all sources of information that contribute to hazard identification and the overall strength-of-evidence warranted; (2) identify sources of divergence resulting from different perspectives and opinions of individuals, thus contributing to conflict resolution; (3) help to identify areas for research and assess their potential impact on decision outcome(s) or their level of confidence. The overall objective of this project is to develop a logical assemblage of the diverse factors contributing to the decision making process to accomplish these three goals.

2. INTRODUCTION

Identifying and assessing sources of uncertainty (levels of support) is essential to evaluation of risk analysis for decision making. Numerous authors and workshops have raised issues relevant to this topic, but there remains a need to knit together: (1) the diverse components of information that may be useful, but variable in availability and quality from chemical-to-chemical; (2) the process of inference from the available knowledge base, including the role of primary data (animal and epidemiological data) and supporting informational sources; and (3) human knowledge and judgment. Identification of sources of uncertainty is necessary but not sufficient. The knitted components described above must provide a fabric of rationality, based on completeness and coherence of assembling all available evidence into a risk analysis.

What is rationality? Can scientists of different disciplines, administrators, decision makers, industrialists potentially affected economically by regulatory decisions, health advisory and other special interest groups, persons whose thoughts and perceptions may be molded by

different backgrounds and interests, share a common conception of rationality? Are there alternative conceptions of rationality? If so, are they compatible within the risk analysis context? Is there a flexible conception of rationality that includes the others as specific cases? These questions are fundamental to our overall objective.

Five fundamental rational bases for decision making need to be addressed before a complete methodology for hazard identification can be created. These questions are:

(1) What does it mean to claim that a risk analysis is rational?

(2) Where would one look in a risk analysis to determine if a claim to rationality was valid?

(3) What are the various views on rationality which might be adopted by competing camps in a debate on risk, and how do these views fall out as specific cases of a more general understanding of rationality?

(4) What kinds of information might appear in a risk analysis and what role does this information play?

(5) What kinds of judgments must the analyst make concerning this information?

Part I of this report consists of several subtopics. Decisions are made by humans and thus reflect human judgment. What is considered as a rational decision depends on many factors, such as objectives (ends), alternatives (means), and beliefs (which include beliefs about nature and about human values). Human judgment, however, with whatever background factors and perspectives influence it, is the means by which decisions are made. No formula for rational behavior is possible that could circumvent this element of human judgment. Furthermore, rationality is a human conception with many variations, as apparent from the long history of philosophical inquiry into that topic. This history of thought on rationality will serve

to provide enlightenment on the breadth and depth of conception of rationality from which decision making for risk analysis may be approached.

"Rationality" is not something that can be discovered and labeled, like a chemical element. It must, instead, be defined consistent with principles of internal coherence and of coherence with other words in the vocabulary. Consequently, as occurs in any attempt to discuss any normative word (concept), one must sometimes tolerate circularity as experienced when one word is described in terms of a second, leading to a third, etc., until the circle is completed to referencing the original word. To illustrate, suppose we ask what minimal characteristics should be included in a conception of rationality to which most people would agree, e.g., logical, reasonable, making "common sense," all of which are well understood terms in frequent use. One is immediately led to further questions of refinement, e.g., logical in what sense? Does it follow deductively from some axiomatic truths? Does it attain a specific goal or objective? Can two different opinions (decisions) based on the same evidence both be logical? Similarly, one could replace "logical" with "reasonable" or "makes sense," or "rational" and the same outcome would pertain.

A useful way for the reader to approach material in this report, perhaps ironically, is simply to think of rationality initially in terms that are comfortable and meaningful, e.g., "logical," "reasonable," etc. This provides a useful starting point from which to expand and perhaps modify one's perspective. The next higher stepping stone into this subject matter is the "Seven Desiderata of Rationality" of Bunge, given in the next chapter under "constructing a complete vision of rationality." Bunge describes seven factors that any claim to rationality must consider, and these factors provide a useful framework for subsequent discussions in this report. The next section (3.2) considers what human skills enhance rational decision making. The intimate association of rationality and human judgment is also fundamental to the topic of the

next chapter. To see this, one need only observe that the "meaning" of rationality is based on human judgment, as is the assessment of the rationality of an action or decision. Twenty areas in which an analyst or policy maker must make distinct judgments that reflect human values are described. Human judgments may differ on assessment of an issue, or on the criteria for assessment, making the relevance of the yardstick subjective. Just as people can often benefit from the opinions of others, risk analysts and decision makers may benefit from the expressions of diverse rationalities. Ten broad classes of rationality that may underlay societal debates, i.e., may produce alternative but rational differences of opinion (judgment), are given to conclude Chapter 3.

Chapter 4 addresses some central issues regarding alternative perspectives on rationality. Whether rationality is viewed as a descriptive or a normative (prescriptive) theory of human behavior may depend on one's discipline (Section 4.1.) A risk analysis must be structured to yield predictions related closely to an existing means for reaching a clear set of ends. This issue is discussed under the rationality of beliefs, means, and ends, the subject of Section 4.4. The most complete vision of rationality incorporates beliefs, means, and ends, judged both intrinsically and through their relationship, which is significant to the risk assessment process. These topics are addressed in Sections 4.3 and 4.4, and summarized in relation to other chapter topics in Section 4.5.

We turn to the history of philosophy (particularly the history of science and analysis) in Chapter 5 to discuss the manner and degree to which beliefs, means, and ends impact on each other. Differences in perspective on the possibility of a logic of science lead to competing views on the extent to which a risk analysis can be driven entirely by the internal goals of science. Classical, empiricist, and rationalist rationality, logical positivism, rational skepticism, and probabilistic rationality are compared and contrasted, and illustrated in the context of risk

analysis. Several issues related to this topic also are addressed to complete the discussion of rationality and beliefs as a basis for risk analysis and decision making.

Throughout Part I of this report, the example of a risk analysis for a carcinogen is employed. In particular, the chosen example is radon in drinking water. This example was selected not because it necessarily is more enlightening, but because one of the investigators (Crawford-Brown) has been involved for several years in efforts by the U.S. EPA Office of Drinking Water[7], the Science Advisory Board, and the American Water Works Association to reach agreement on a regulatory standard. The example is not developed in full detail, since that detail is the subject of future research. It is intended, however, that the present report summarize all of the considerations which will form the basis for that detail. As mentioned earlier, reviewers are asked to read the present report for inconsistencies, missing principles of reason, poorly or incorrectly defined terms, and (of most importance) a lack of clear relevance to the *specific* issues which invariably arise in pushing a risk analysis through from conception to justification.

Some of the sections of this part of the report contain fairly detailed analyses of philosophical concepts. Readers may find it useful to skip those sections initially, focusing on the summary sections to gain a broad overview of the kinds of issues raised and potential solutions developed. A second reading might then be used to gain a more detailed understanding of how the summary information arose.

Rationality is addressed in this project on uncertainty in risk analysis to provide an informed basis for identifying needed sources of information, processing and evaluating it for credibility and relevance, and then making rational judgments. What information is needed for risk analysis? Which sources are preferable? Which are of greater significance for drawing conclusions? What assumptions are implied by "default" guidelines when information is

unavailable? How should expert opinions be weighed vis-a-vis experimental data? There are many facets to the topic of uncertainty in risk analysis. Consequently, an approach that is systematic and developmental, i.e., that can be expanded, revised, and refined as needed, may be the most tractable approach. The basic components of such an approach are described briefly in this part of the report.

3. CONSTRUCTING A COMPLETE VISION OF RATIONALITY

Decisions are made by people, some of whom are more skilled at making rational decisions than others. What are these skills, i.e., what characteristics are found in a good decision maker? This topic is treated in Section 3.2, following an introduction to the topic of rationality via Bunge's "Seven Desiderata of Rationality." The discussion of Bunge's principles in Section 3.1 is a useful departure into the topic because he describes characteristics that any rational system must have, using common terminology. Having described the features of rational decision making and the personal skills that enhance its implementation, twenty subject areas are given in which the analyst/decision maker is called-on to make judgments that reflect human values (Section 3.3). Disputes and disagreements will still arise in risk analysis, however, which leads us to consider alternative classifications (or types) of rationality that may be encountered in society, e.g., supporting adversarial positions (Section 3.4).

3.1. Seven Features Essential to a Rational System

Bunge[45] describes seven characteristics essential to rationality which he calls the "seven Desiderata of Rationality." they are included here as a framework for subsequent discussion.

1. Conceptual--a properly rational system minimizes fuzziness, vagueness and/or ambiguity in its terms. The importance of this feature was highlighted by Wittgenstein[46], who felt that logic could not begin until language itself was placed onto a well defined basis. All people using a term must mean the same thing,

and this meaning must be chosen to reflect accurately the objective features of the world. In his early work, Wittgenstein argued that the proper definition of a word could be obtained through careful analysis of objective reality. His later work[47], however, emphasized the role of social agreement in choosing the meaning of words. Regardless of the viewpoint chosen or the origin of a word's meaning, rationality involves an attempt for conceptual clarity which allows a comparison of the beliefs of individuals through a shared language.

2. Logical--a properly rational system strives for consistency and a lack of contradiction. Beliefs (means, ends) are, to the degree possible, obtained through the application of well defined rules of reason. This focus on logic is intended to ensure that beliefs (means, ends) can be shown to be reasoned and not the result of judgments which might be made as matters of convenience. As Nathanson[48] writes, "Rationality, then, involves a striving to be objective, and objectivity involves the attempt to discount those features of ourselves or our situation that might involve our judgment but that are not relevant as evidence." The rules of logic supposedly open choices to public discourse.

3. Methodological--a properly rational system produces a habit of questioning beliefs (such as in Popper) and/or verifying beliefs (as in the Vienna Circle writers such as Carnap[49]). Without the presence of these methodological rules, people would have no shared approach to rooting out and changing false beliefs.

4. Epistemological--a properly rational system accepts evidence only when that evidence satisfies criteria of quality (such as empirical accuracy, relevance, etc.). The epistemology instills in the rational person a love for particular kinds of evidence (such as the logical positivist's desire for observation). Statements which are pure conjectures, unsupported by "appropriate" evidence, are avoided in the process of reasoning.

5. Ontological--a properly rational system uses only terms (such as force, cell, etc.) which are believed to be descriptions of the world. Since Bunge believes strongly in science, his own ontological criterion is that all terms used in reasoning must be those used by science and technology. Others may disagree with him, but all rational people will apply reason only to terms which describe entities and relationships believed to exist in the world.

6. Valuational--a properly rational system strives for goals which have been determined to be worth attaining. These goals should, moreover, be the highest goals.

7. Practical--a properly rational system adopts means for action likely to yield the desired ends. People may disagree as to what is to be meant by likely. In some approaches, all of the available means are described in detail and the evidence for their ability to yield the desired ends weighed. The means most likely to reach those ends from amongst the competitors is selected. In other approaches, only means which satisfy some minimal level of crafting form a rational basis for action. Beliefs that fail to satisfy this minimal level are rejected even if they are

the best amongst the competitors. In both approaches, however, the key to practical rationality lies in demonstrating a clear link between means and ends.

Bunge moves from these seven desiderata to a distinction between various levels of rationality. People who rely on only a subset of the seven are said to be *semirational*, a wording similar to Simon's idea of bounded rationality.[50] People who rely on all seven are said to be *fully rational*. Anyone who rejects all of the criteria is said to be *non-rational*. Only a person who accepts the criteria but goes deliberately against the outcome of rational analysis is said to be *irrational*.

3.2. Skills Required by the Rational Risk Analyst

What skills enhance rationality? Rescher[51] has identified a need for five distinct faculties which must be brought to bear in rational discourse or action. There is a need for imagination, since the rational person strives to look for alternative beliefs, means, and ends to be judged. A firm grasp of information processing is required in order to form beliefs. Skills at evaluation of alternatives (beliefs, means, ends) give the capacity to weigh the merits of those alternatives. The rational person must also be able to select an option through informed choice. To some views of rationality, such as the classical theory, selection is dictated by the methodology of rationality. In other views, such as that by Berkson[52], rational inquiry is a guide without determining the final choice between options. In Berkson's words: "The basic idea is that the individual, not the method, makes the choice; but the individual should be influenced in his decision making process by rational argument." Finally, a rational person requires the skill of agency, the capacity to actually carry out a choice.

3.3. Value Judgments in Rational Action

The preceding discussion suggests twenty areas in which an analyst or policy maker must make distinct judgments which reflect human values. Some of these areas will be clear only after a reading of Chapter 5, which includes a discussion of the failure of philosophy to locate a complete logic of science and evidential reason. These value judgments are listed here in no particular order. Each judgment will be subject to dispute and complete rationality requires that each be discussed in a setting which includes all affected parties, as required by game rationality (see Section 3.4). At the highest level of judgment, it might also be asked how the specific conceptions of rationality listed in Section 3.4 should be valued. The example of regulating radon in drinking water is used throughout the following discussion.

(1) Relevance of evidence to the stated problem--evidence may be assembled from a potentially infinite number of areas of human study. A radon risk analyst must judge the relevance of data from experiments on the breakage of DNA by radiation, from the transformation of cells in vitro, from small animal studies and from human epidemiological studies. There also are experiences suggesting the general reliability of the scientific community at transmitting results to various groups in society. All of this evidence must be judged for its relevance to attempts at predicting the effects of radon in the home.

(2) Epistemological status of evidence and beliefs--some people may look for direct empirical evidence that removing radon from water has resulted in lowered risks *in practice*, i.e., in historical cases. Others will be satisfied by deductions from theories which predict the desired effects in the absence of any human experience with mitigating radon. Still others may be satisfied by extrapolation from findings at high levels of human exposure, using either theory-free curve fits to the existing data or explicit theories about the role of radon in producing or reducing

cancer. A judgment must be made as to the weight assigned to these varying kinds of evidence and how that weight justifies particular courses of action such as public warnings, expenditures for mitigation or the imposition of regulations.

(3) Level of causality--all phenomena such as health have many levels at which causality may be assigned. In the case of radon, the cause of cancer may be said to be the deposition of radiation energy in cells. It may be said to be the ingestion or inhalation of radon. It may be said to be the presence of radon in the water. It may also be said to be the political system which allows such levels, or the economic system which allows people to become poor, undernourished and, hence, susceptible to the effects of radiation. Each of these levels of causality suggests a different strategy for mitigating risks. A judgment must be made of the appropriate level at which causality will be analyzed.

(4) Clarity of terms required--some ideas, such as long-term frequency, are highly technical and well defined. Others, such as human confidence, are less well defined but potentially incorporate a wider range of considerations. A judgment must be made of the clarity required before a term or idea qualifies as a valid basis for analysis.

(5) Degree of bounding--no risk analysis can include all of the factors related to even a single end. If radon is removed from drinking water by aeration, it will enter the air. If it is removed by absorption onto charcoal, it will enter the soil of a landfill. From there, it may migrate to other parts of the environment. Given the complexity of environmental systems, only a select sample of the potential pathways of exposure can be considered. The model of risk must, therefore, be

bounded. A judgment must be made as to where the boundaries are to be imposed in light of the stated goal to protect the public health.

(6) Specification of ends--the EPA desires to protect the public health. This end may, however, include minimization of the average chance of fatal cancer, minimization of effects in children, minimization of years of life lost, etc. In addition, there may be the desired ends of lowest cost, degree of democratization of choices, etc. A judgment must be made of the most significant ends and of how the process of analysis will reflect those ends.

(7) Dealing with uncertainty and ignorance--a problem may be viewed as a choice between existing options. It might be argued that the risk analyst must choose between existing theories for predicting the effects of radon on health. The analyst then is in a state of uncertainty. It might also be argued that none of the existing theories is well established, requiring an admission of ignorance. A judgment must be made as to whether an analysis of uncertainties is required, how that analysis should affect decisions, and whether the possibility of ignorance should be factored into the decision in addition to uncertainty.

(8) The will to believe specific predictions--as James has pointed out[57], most questions cannot be answered with certainty. If the question is not of great importance, it may be acceptable to withhold belief until the evidence improves. If the question is of great importance, however, it may be necessary to adopt a belief despite great uncertainty and ignorance. This new belief might concern a particular prediction of the effect of radon, or it might concern the more vague belief that the effect lies within certain bounds (such as the belief that the chance of cancer associated with a concentration of X is less than or equal to Y). A

judgment must be made as to whether belief is warranted and what kind of belief is warranted in light of the existing state of evidence.

(9) Sufficiency of logical necessity--no beliefs follow necessarily from a firm foundation of truths (see Chapter 5). And yet some beliefs are more firmly rooted in observation and logic than others. There is a natural desire to see beliefs reduced to logic, since this provides the easiest route to public scrutiny of the link between evidence and belief, and avoids the need for judgment. But logic ignores skills of crafting and intuitive insights gained from experience (typically called engineering judgment). Toulmin[58] conceives of the process of reasoning as one of offering warrants for belief. Warrants do not guarantee the truth of a conclusion and, therefore, do not lead to a strict logic. Yet they do count as a kind of support for any conclusions. A judgment must be made of the degree to which predictions of the effects of removing radon must follow the rules of logic and the degree to which it may require the less stringent idea of warranting. A judgment also must be made as to whether a particular warrant is sufficiently weak to require strong public scrutiny or sufficiently strong to justify acceptance without stringent requirements for debate.

(10) Location of rationality--a judgment must be made as to whether an action is rational if either the beliefs, means, ends, or the relationship between these parts may be left unexamined in a claim to rationality (see Section 4.2).

(11) Role of conceptual and empirical success--as mentioned in the discussion on Laudan[36] (see Chapter 5), theories may be assumed both to yield predictions of the outcomes of experiments and to provide explanations which unite concepts within a field of science. The former is a requirement of foundational truth and

the latter is a requirement of coherence. In the case of theories concerning the effects of radiation, some give explanations referring to a wide range of biological factors (DNA damage, cell replication, etc.) while other refer only to physically unspecified "thresholds" for cancer. Even if both predict the existing human data equally, a judgment must be made of the importance of the former's conceptual sophistication, its ability to unite biochemical, in vitro transformation, and epidemiological data.

(12) Selection of humans--if writers such as Longino[37] and Polanyi[40] (see Chapter 5) are correct, all rational activities, including science, require distinct judgments and tacit skills. A question arises, therefore, as to how individuals possessing such skills are to be located. Two features of the selected humans seem to be essential. The first is an adequate base of experience from which skills may develop. In the case of radon, this may require selection of scientists in the field of radiation biophysics. The second is an adequate degree of reflection on the quality of that experience and the resulting skills. This may require selection of people familiar with rationality or legal argument. In addition, game rationality requires that a range of humans with (perhaps unstated) differences in ends and skills be incorporated into the process of decision.

(13) Completeness of the "hard look"--with limited time and resources to devote to an activity such as the analysis and mitigation of radon risk, it may not be prudent to suspend belief awaiting collection of all available data. At some point, it must be judged that both the pool of data and the extent of critical reflection is adequate for decisions. This requires a judgment of when an analysis constitutes a sufficiently "hard look" to act as a rational basis for a particular course of action.

If the course of action is publication of a pamphlet warning of radon risks, the judgment of sufficiency may be different than would be true for setting regulatory limits.

(14) "Knowing that" versus "knowing how"--two judgments will be required in this regard (see Chapter 5). It is necessary to determine when a claim to understanding, unsupported by a demonstration of the efficacy of a means in practice, is an adequate base for rational action (such as may result if predictions of the effect of environmental levels of radon have not been tested through tests of mitigation). It must also be determined when a claim to limited success in practice may form the basis for rational action if an understanding of that success is not available.

(15) Interlocking of beliefs, means and ends--in an ideal case unconstrained by contextual rationality (see Section 3.4), all beliefs, means, and ends might be examined. In practice, however, only beliefs relevant to acceptable means, or means relevant to acceptable ends, will be considered. For example, while changes in nutritional status might aid in lowering susceptibility to the effects of radon, the means for achieving this might lie outside those allowed to the EPA. Similarly, only cancer risks typically are considered in specifying ends. A judgment is necessary concerning the degree to which these three areas should be explored completely and in isolation from each other, at least prior to the time for decision.

(16) Extent of coherence--as proposed by Quine[44], rationality implies a coherence between the separate beliefs within a system of belief. It is not clear, however, whether all beliefs must cohere when reaching for specific ends. Should, for

example, religious beliefs be allowed to require a well developed consideration of the influences of sin on the risk from radon? Should a belief in benign nature modify the analysis? Must a theory concerning the risk from radon include explicit consideration of oncogenes?

(17) Ontology--one criterion of rationality was that the system of belief employ entities and relationships (such as cells, transformations, etc.) which are judged to exist in the world. Some concepts such as states of emotion, confidence, psychological causes, etc., may fall below a predetermined ontological status. There is a debate in the modeling of health effects over whether the stages of a multistage model of carcinogenesis satisfy this criterion. A judgment must be made of the factors which will be allowed in any explanatory and/or predictive theory.

(18) Choice of fides--if Fideism (see Section 3.4) is chosen as the basis for rational action, a judgment must be made as to which beliefs are no longer open to rational debate. These might, for instance, include the belief that a cancer may be identified without question. Other beliefs, such as that particular patterns of scattered light indicated DNA damage, may not satisfy the judgment of sufficient belief to constitute a fide. They will require additional support from more well established bodies of evidence.

(19) Ambiguous evidence--for all beliefs, there will be evidence in favor of, against, or neutral to, that belief. The ability of radiation to produce DNA breaks acts as support for the belief that radiation is an initiator. The cytotoxicity of radiation supports the belief that it may be a promoter. A judgment must be made as to

how these various instances of partial confirmation and partial falsification should be weighed into an overall assessment of belief.

(20) Concept of the weight of evidence--the degree of belief in light of evidence may be given by a purely subjective statement of confidence, a statement from sampling theory, a statement deduced from a Bayesian perspective, etc. This weight may be numerical or verbal, and may or may not refer to psychological states. A judgment must be made as to how the concepts of weight of evidence, probability, confidence, etc., are to be interwoven and defined.

It should be clear that a large number of judgments potentially appear in any claim to rationality, even if attention is focused on the rationality of belief. All of these, in fact, underlie such claims even if the judgments are not made explicit. Any framework for rational analysis, decision and action should, therefore, make the above judgments explicit and the subject of debate. It is important to bear in mind that the particular values brought to the forefront in making those judgments may depend upon the activity in which one is engaged. An ordering or assessment of values applied to the rationality of scientific research may be inappropriate for regulatory action, given the potential for different ends. No attempt has been made here to specify the values which should be adopted. The intent, instead, has been to reveal their existence and to place them within a framework of rational thought and action.

3.4. Broad Classes of Rationality Underlying Societal Debates

This section closes with an attempt to catalogue various types of rationality which might be found in society. Any person or group might employ more than one type of rationality in forming beliefs or in carrying out a decision or action. Still, it is useful to establish a typology to aid in locating the fundamental assumptions adopted by that person or group when a claim to

rationality is made. The following summaries, then, represent a rough sketch of the various stances which might be adopted in the search for, or in construction of, a rational system. An example using the case of radon is given at the end of this section.

1. Classical rationality--this is the belief that rationality essentially is equal to truth, and truth is obtained through the use of logic (see Chapter 5). This logic is applied to a set of premises about the world which are taken to be absolutely certain. The classical camp might usefully be divided into two schools with different epistemologies. The empiricist school (Locke, the logical positivists) choose their firm premises or foundations from the results of observation. The rationalists (such as Descartes) chose their firm premises from "clearly and distinctly perceived" items of introspection. In both cases, only beliefs following deductively from these foundations are allowed in the quest for rationality.

2. Process rationality--this is the belief that no statement or action will satisfy the desires of classical rationality. All beliefs are subject to question since there are no firm foundations. Rationality implies, instead, a process of constant questioning and a search for more evidence. An example is Popper's idea of critical discourse with repeated attempts to falsify a theory (see Chapter 5). It is the process, and not the current state of match between evidence and belief, which justifies a claim to rationality.

3. Fideism--a "fide" is a faith, something taken without complete support in evidence. The Fideist's (such as Polanyi) admit that the classical ideal cannot be met since there is no completely firm foundation for belief. They argue, however, that some beliefs (particularly those of well-tested science) are sufficiently strong to justify acceptance as a matter of faith. People will, of course, disagree as to which beliefs satisfy this criterion of being a proper fide. Once the fide is chosen, however, this approach is similar to the classical approach to rationality.

4. Probabilistic rationality--the foundational assumptions required by the classical approach may be said to be somewhat less than firm. They may, however, be given a degree of support through probabilistic ideas. These probabilities may be said to be an objective property of the world or of our methods for studying the world (as in long-term frequency approaches to sampling theory), or a measure of our state of confidence. As with Fideism, rationality then is obtained in a manner similar to the classical approach once the probabilities are assigned to the probabilistic premises. To be rational means to act on the belief possessing the highest probability of being true.

5. Limited rationality--even scientific models fail to predict the entire functioning of the world. Causal factors are left out because they are not understood. Others are left out for computational convenience. It is not possible, therefore, to reach all of the possible ends through the use of such models. Still, a rational person may reach a few well prescribed ends through the use of bounded reason. There

will, of course, be disagreement as to what constitutes proper and acceptable bounds on the ends, means, and beliefs appearing in a rational discussion.

6. Contextual rationality--complete rationality might require a very large amount of time to study a problem, particularly if the problem is complex. Humans, however, have competing ends which require attention. It is important, therefore, to allocate time so that a single problem does not attract an inordinate amount of attention. A rational person analyzes a problem only to the degree feasible in light of other demands on time. The goal is to limit analysis of specific problems so that the full range of problems may be approached in at least an approximate manner.

7. Game rationality--the classical theory assumes that all of the ends and methods for judging means or beliefs can be ranked from highest to lowest (or best to worst). While this still may prove to be the case, agreement on the appropriate ranking has not been found in society. As a result, any societal debate will be characterized by different groups adopting a different ranking. The rational person treats this situation like a game in which the competing views must be balanced through a process of interaction. In this sense, the view bears some resemblance to process rationality (#2).

8. Adaptive rationality--it is a common feature of human understanding that beliefs change with new evidence. A properly rational person, therefore, looks for a process (implying a kind of process rationality) which mimics the growth of knowledge. An example of this approach is Hesse's "learning machine" model of science[53], or the Bayesian methodology of updating beliefs in light of new evidence.[54]

9. Selective rationality--if, as Longino emphasizes[37], human judgment is essential to assigning relevance to evidence (see Chapter 5), it would be best to let those judgments be made by the most qualified people. Selective rationality then involves the identification of groups or individuals with the best skills at judgment. This is by no means a simple matter, since it is not clear how those skills are to be identified. The problem is particularly difficult when ends, means, and beliefs become inseparable, since no individual or group is likely to possess skills in all three areas of judgment.

10. Posterior rationality--the classical theory, as well as modern decision theory, assumes that the ends are known before rational analysis begins. The means to reach those ends then are assembled, beliefs assigned to those means, and the appropriate means selected. At times, however, the ends may not be known prior to analysis. They may, instead, be a product of an analysis, such as when scientific research discovers the importance of species diversity, which becomes a new end. A rational process should, therefore, be capable of generating new ends or of changing the direction of analysis once new ends are identified in the process or reasoning.

These various conceptions of rationality can be illustrated by considering how they might cause a person to view a risk analysis for radon. It should be emphasized that the conceptions are not necessarily mutually exclusive. There should be a healthy regard given to each conception, since each provides valuable insights into a complete picture of rationality.

Assume that the EPA wishes to regulate radon in air or water, and that a claim to rationality is desired. The classical rationalist will focus attention on the set of beliefs concerning how radon is related to a given end (here the production of cancer). Only predictions which must follow logically from observations (empiricism) or "clearly and distinctly perceived" statements (rationalism) will count as being true. Only these beliefs, therefore, will count as rational products of the risk analysis. All other beliefs, regardless of their status as being partially confirmed, will be rejected as non-rational. While the classical approach certainly stands as an ideal, it is a simple matter to find potential flaws in any prediction concerning radon and, hence, the classical skeptic could argue that all risk analyses were non-rational.

The fideist would search for some principles which could be taken as reasonable in beginning the process of reasoning. This analyst might argue that an observation of an excess incidence in an exposed population was a reasonable basis for belief, even if (as Hume argues) it does not certify that belief. The analyst might also argue that a linear dose-response function has been reasonably well established. Such fides then could be used to analyze the existing data in analytic fashion. Given the fides, all of the predictions of the risk analysis would satisfy the classical ideal. Different analysts probably would argue as to the appropriate fides, but at least it would be clear what the argument concerned. Only beliefs following deductively from the fides would be counted as rational products of the risk analysis.

The probabilistic rationalist would argue that it is incorrect to adopt any principle (such as a linear model) as a fide. Instead, all assumptions must be given a probability of being true.

Deductions from the assumptions would be developed and the calculus of probabilities used to assign a probability to each deduction. The rational result of a risk analysis would be a set of predictions concerning the effect of radon on lung cancer, the assumptions underlying those predictions, and the probability assigned to each prediction. These probabilities might be obtained either through statistical considerations or by specification of a state of confidence. In either case, the search is for the prediction (of cancer attributable to radon) which possesses the highest probability.

The above approaches focus on the quality of beliefs about the link between radon and lung cancer. A second party might argue that morbidity is equally as important as mortality. He might also point out that genetic makeup predisposes some people to lung cancer. This would lead him to claim that the risk analysis by the first party was not rational since it lacked consideration of morbidity and genetic predisposition. The first party would counter that the EPA said to consider the average risk of lung cancer, and that while this is a limited goal, the analysis still has limited rationality. The two parties disagree as to when the limitation is so severe as to require withdrawal of the claim to rationality.

While this debate was going on, the first party might point to other concerns in society. There are needs other than the control of radon. Pesticides, food additives, etc., must also be regulated. Increasing the considerations taken into account in analyzing the risk of radon would draw resources away from other analyses. This might increase the lives saved from radon mitigation, but at the expense of ignoring other risks in the interim. An appeal is being made here to contextual rationality. The limited rationality of the bounded study (which considers only the average risk of lung cancer) is justified by referring to the larger societal context within which the analysis occurs.

The risk analysis for radon also presumes a set of ends. It might be assumed that the end is to regulate radon effectively, or to do it at the least cost, or to set up an atmosphere within which regulation itself is given historical precedent. The focus might be on lung cancer or pneumonia, on old age groups or young. Different groups usually will not agree as to the appropriate goal in dealing with radon. Game rationality requires that this difference be acknowledged and a process of analysis put in place which allows the groups to interact. This can be accomplished by opening the process of analysis to scrutiny, making it clear how different assumptions concerning the goals are reflected in different results of the analysis.

The analyst would need to make judgments concerning the "quality" of beliefs generated in the analysis. A hope of the classical rationalist is that this quality can be judged on the basis of a logical relationship between the evidence and the belief. Different groups of researchers would weight the evidence differently. For example, epidemiologists favor even poorly controlled human studies in uranium mines over well controlled animal studies involving exposure of rats and beagles in radon chambers, while the reverse judgment is made by those conducting the animal experiments. The rationality of the final risk analysis depends heavily, then, on the proper selection of experts chosen to give evidential weight. The claim to rationality would rest on the ability to demonstrate that the analysts were chosen for appropriate skills.

As a risk analysis for radon proceeds, new perspectives will develop. Information deemed irrelevant prior to the study will gain relevance through the analysis. No predictive models of the effects of radon might be developed, requiring a reassembling of relevant data. This might require drawing upon experts not identified in originally structuring the analysis. A procedure must be put in place to ensure that the boundaries of the analysis are not restricted to those established prior to the analysis. This procedure, then, justifies the claim to adaptive

rationality. Similarly, the analysis of radon might reveal that morbidity is a serious outcome deserving of attention, despite a prior commitment to examining only lung cancer mortality. The control of morbidity would constitute a new goal of the analysis. Procedures put in place to ensure that the risk analysis was capable of generating new goals would satisfy the claim to posterior rationality.

4. WHERE IS RATIONALITY FOUND?

The following section provides a sketch of the way in which people might differ in speaking about rationality, at least to the degree that they differ over the aspect of the analysis which might be called rational or irrational. A key feature to be noted is the possibility that the rationality of a risk analysis cannot be judged without also understanding the purpose behind the analysis (i.e., what the analysis is to be used for).

4.1. Descriptive and Prescriptive Rationality

One of the central topics in rationality in areas such as decision theory is whether it is a descriptive or normative (prescriptive) theory of human behavior. Psychologists and behaviorists[8,9] tend to view rationality as a scientifically describable and discoverable process which is found in actual human behavior or thinking. Their goal, then, is to demonstrate the dynamics by which specific humans or groups come to be rational, irrational, or non-rational.

The intent is descriptive and/or predictive, and their approach may be used to help discover the underlying goals and cognitive processes leading to specific risk analyses or decisions.[10,11]

Contrasting with this approach is the school of thought which links rationality to an ideal of human reasoning and/or behavior, which may or may not reflect actual historical cases.

Rationality, then, is a capacity of humans which must be defined and nourished. Lying primarily in the realm of philosophy, this school searches for an ideal conception of rationality which will be used as a yardstick for measuring the quality of human thought or behavior. The focus is not on how people *actually* do risk analyses, but on how they *ought* to do them.

Both of these views are found necessary in the present paper. If an individual or organization is to resolve disputes, or at least keep them under some measure of control, it is necessary to understand the rationality of specific disputants. At the same time, it must be admitted that people rarely meet an ideal of behavior, however well intentioned they might be. The problem, then, is to design a systematic way of thinking which will help people become rational or to improve on their rationality. Descriptive theories describe society. Prescriptive theories contribute to the evolution of society. Both theories are needed in determining the reasons for disputes and in establishing a framework for judging the relative merits of competing positions in the dispute, or for resolving the dispute.

4.2. The Rationality of Beliefs, Means and Ends

Putting aside the issue of description and prescription, it is then necessary to find what is being described and/or prescribed in discussions of rationality. What would it mean to suggest that an EPA regulation limiting the concentration of radon in drinking water to level X (such as 300 pCi/ℓ) is rational? To answer this, it is necessary to isolate the components of such a regulatory decision. In the most general sense, it might be argued that EPA:

(1) has a goal of producing a world in which the chance of death from cancer caused by radon is Y (where Y might be 10^{-6} or 10^{-4});

(2) believes that a concentration of radon equal to X produces such a world; and

(3) chooses to limit the concentration of radon to X.

The three components to this process of reasoning must be subject to the constraints of rationality. It might be asked if the *belief* of the Agency (that a concentration of radon in water equal to X will yield a chance of death Y) is rational. In this case, the intent is to show that the belief follows in some reasonable sense from evidence about the nature of radon and its causal link to cancer. To be rational, then, requires that the various beliefs available to the Agency be subjected to scrutiny to determine their degree of support in the evidence. This situation will be referred to as the *rationality of belief.*

By the same token, the decision to limit concentrations to X is only rational because the goal (or *end*) is determined to be a world where the chance of dying is Y. This rationality thus requires a reasonable Agency goal. Some other group, such as a water treatment association, might argue that the proper goal is to minimize expenditures on water treatment. The EPA might respond by stating that its primary (or highest) goal is protection of public health, and that economic factors are less important. This examination of the *ends* of action will be referred to as *the rationality of ends.*

Finally, the Agency must select a particular *means* to reach its stated ends. The means might, for example, be to limit concentrations in water to X, alter the use of water supplies, provide warnings, etc. Presumably, these means will be consistent with both the beliefs of the Agency and the ends. It is assumed that the means will lead to the ends in a demonstrable manner. Choosing means appropriate to the ends will be referred to here as the *rationality of means.*

This suggests three areas in which the rationality of an Agency decision or action might be challenged. Discussions might focus on the various beliefs available to the Agency, such as whether 100 pCi/ℓ, 300 pCi/ℓ or 10,000 pCi/ℓ yields a chance of death equal to Y. Rationality then would require that each belief be supported by evidence concerning the fundamental

physical laws governing radon and cancer. A choice must be made as to how this belief might be characterized. In the field of risk, this step typically is though of as *risk analysis*.

Discussion might also focus on the stated ends, requiring a clear demonstration that a world containing a chance of death from radon exposure equal to Y should be the ultimate goal. Others might argue that the proper ultimate goal should be something else, e.g., human happiness, and that the chance of death is only one factor in assessing this goal. They would demand proof that reducing the chance of death to Y without consideration of cost truly leads to the state of greatest human happiness. Their judgment of the rationality of an analysis would be determined by assessing whether that analysis considered aspects of risk related to the most significant ends. In the field of risk, this concern is loosely termed *risk policy*.

The discussion also might center around the appropriate means to reach the goal. The EPA might argue that aeration of water is preferred to other means such as removal by activated carbon.[12] A water works association might argue that the EPA should first identify individuals sensitive to radiation, and then limit concentrations only in their homes. The disagreement here revolves around the most efficient means to reach a given end, conditional on a specific belief about the world. The rationality of the analysis would be judged according to the ability of the analysis to provide appropriate solutions. This area of risk usually is referred to as *risk mitigation*.

Controversies in the philosophy of rationality focus on the relative importance of these three aspects of an action in claims to rationality, and on the manner in which the three are interrelated. Logicians and most "pure" scientists tend to focus on the state of belief. To be rational means to base actions on beliefs which satisfy some criterion such as "minimal confidence" or "being the best available belief." The rationality of a risk analysis could then be judged without reference to any uses of the analysis. Technologists, "applied" scientists and

policy analysts tend to focus attention on means. Their criterion for rationality is that the most efficient means be adopted for reaching a given goal. Only analyses which focused on aspects of risk related to known techniques of mitigation would count as being rational. Ethicists and policy makers tend to focus attention on the stated ends. Their criterion of rationality is that the chosen ends in some sense be the "highest" ends, the most noble expression of human needs and desires. Only analyses which focused on those aspects of risk associated with the highest ends of society would count as rational.

These three groups can disagree fundamentally about the ability to be rational in all three senses. Logicians with a leaning towards empiricism tend to argue that ends have no rationale. Taking their cue from David Hume[13], they think of ends as being mere matters of taste bereft of any reasonable support, and of means as being matters of crafting. To them there is no sense in which either ends or means are the subject of reasoned debate. In this classical theory of rationality (discussed in more detail in Chapter 5), rationality is equated with truth, and truth consists of a perfect belief about reality. Evidence is gathered and shown to support a belief, such as that a concentration of radon in water equal to X yields a chance of fatal cancer equal to Y. The most rational belief is the one corresponding most closely to the features of the world. Such people do not argue that means and ends are not important. They simply claim that these aspects are matters of unreflected skills and tastes, respectively. The rationality of dealing with risks is located in the choice of an appropriate predictive belief, assumed to be the best scientific estimate.

Those concerned with the rationality of ends tend to be less concerned with the logical relationship between evidence and beliefs. This is not to say that they deny the need for such a relationship. They see the relationship as being an issue of logic, which they separate from rationality. To be rational is to guide life according to the highest principles, which may include,

but is not restricted to, consideration of the "truths" of science and logic. Evidence and belief are useful only as an aid in reaching a higher goal. This focus is termed "aim-oriented rationalism" by Maxwell[14]. He insists that to be rational it is necessary to ask *why* a particular belief is needed. He speaks of the rationality of belief, with its focus on logic, as leading to "neurotic science." Whereas those committed to logic search for ever greater precision (risk estimates accurate to several decimal places), Maxwell argues that a focus on beliefs is rational only to the degree it allows humanity to reach "higher" goals, e.g., relieving hunger, poverty, and disease. The most rational enterprise begins by examining goals and tailoring the pursuit of belief to satisfy those goals adequately.

Maxwell's approach shades over into the "rationality of means." The history of philosophy prior to the 19th century tended to revolve around debates between rationalists and empiricists. These two schools of thought agreed that beliefs could be placed on a firm "foundation," giving rise to the term foundationalism. Scientific theories were considered true once they rested on such a foundation of established facts. Rational people sought to base their beliefs upon "real" features of the world, which were taken to be objective features unrelated to human conceptual schemes. Rationalists and empiricists disagreed, however, as to how *realism* was to be found. Empiricists relied on experience, particularly sight or observation.[15] Rationalists relied on clearly perceived human insights much as in mathematics. Once these foundations were obtained, all of the features of the world could be predicted. These predictions would be analytic consequences (hence, the word analysis) of the foundational beliefs. The only trick was to ensure that the foundations were real, objective properties of the world.

The views of rationalists and empiricists can be contrasted with "instrumentalism" To the instrumentalist, what matters is human action, not belief. A belief is "true" (and therefore a

rational basis for action) only to the degree that it produces a means for reaching a desired end. A scientific theory may or may not refer to real features of the world. This can never be known; there is no foundation for truth. A theory gains its power instead by allowing humans to *do* something in the world. A theory is nothing more than an instrument and must not be confused with reality. Most scientific theories are, after all, eventually overthrown by a new conception of the world. Nevertheless, the older theories were not considered completely useless. They at least allowed selected actions to take place within limits of error; but they provided limited means for reaching human goals.

Under instrumentalism, a given means to reach an end becomes rational because it "works." The unexplained skill of the scientist or engineer, which embodies what Polanyi[40] calls "tacit knowledge," even can be a rational means for reaching an end. This is true even if it is not understood *why* it reaches that end. In this sense, removing radon from water becomes a rational means of lowering cancer if it has "worked" in the past, even in the absence of understanding how radon produces cancer. As Scriven[16] says, "the proof one knows how to do something is doing it, not talking about it." He is contrasting here the ideas of "knowing that" and "knowing how." To an instrumentalist, the latter world always takes precedence over the former. A rational person must give evidence that a means leads to a specific goal *in practice*. The distinction between "knowing that" and "knowing how" is roughly equivalent to Aristotle's contrasting ideas of theoretical and practical knowledge.[17]

4.3. The Separation of Beliefs, Means, and Ends

In the area of risk, there is a merging of the rationalities of beliefs, means, and ends. It certainly should be hoped that beliefs about the effects of radon on the chance of fatal cancer are as well established as it is possible to obtain. Greater precision of prediction is, therefore,

an important sign of rationality. But there are an infinite number of aspects to a risk. The scientist can study the risk of various kinds of effects, of the interaction between radon and other substances such as environmental tobacco smoke, of the influence of age or nutritional status, of the role of social forces in causing people to use a specific amount of water, and so on. The risk analyst must select what to study from the large number of choices; there are always constraints of time, energy, and cost. This is where ends can modify a risk analysis, influencing the direction of attempts to develop belief. The policy maker does not simply want well established beliefs, which may be the goal of the analyst in a scientific research setting. Rationality requires that these beliefs be researched with particular ends in mind. The relationship may even be reversed. Ends may be adjusted to correspond to areas where beliefs are strong and where there is, therefore, a reasonable expectation that the ends can be reached.

Means also may influence both beliefs and ends. A research project on the risk from radon (i.e., on beliefs about radon) may be stopped well before the risk is fully understood, if the existing understanding is sufficient to allow a selection of means. Scientific research such as an experiment may be viewed as providing evidence for a belief. It may also be viewed, however, as a concrete example of how to modify the world. An experiment in which radon is lowered and the incidence of fatal cancer drops certainly helps in testing theories about how radon yields cancer. But it also provides evidence that physically lowering the concentration of radon is a means of lowering the incidence of cancer. This means is available even if it is not understood why it works. Society may be unconcerned with looking for strong beliefs about radon, i.e., well established theories, if means for reaching ends already are available. Similarly, strong beliefs which are at present useless in suggesting means may be counted as irrelevant in the quest for rationality. It may even be the case that some aspects of the risk, e.g., the role of social forces, may be ignored in the analysis because the policy maker is not willing to consider

mitigation measures based on those aspects. As a result, beliefs, means, and ends subtly interact within the area of risk. This suggests that the relationship between component parts is an aspect of a rational system. This topic will be addressed further in the next section.

4.4. Rationality of Parts and Rationality of Relationships

To complete this section, it will be useful to focus on a separate distinction which might be drawn in thinking about rationality. It might be required that particular beliefs, means, or ends be analyzed independently, such as occurs when risk analysis, mitigation, and policy are addressed separately. Each endeavor focuses on a belief or a means or an end without reference to the other two aspects. Consequently, rationality lies in the belief, or the means, or the end itself, depending on the task. This view will be referred to as *intrinsic rationality* and suggests that the risk analyst, the risk mitigator, and the risk policy maker can perform their tasks without consideration of the others.

Alternatively, it might be claimed that rationality concerns the *relationship* between beliefs, means, and ends. In this case, it is not the beliefs, means, or ends themselves that are questioned. These become the *givens* of life. They may be true, useful, just, etc., but they are not in themselves rational. Rationality requires, instead, a specific relationship between these parts. It requires that society act to get what is desired given that the world is believed to be in a particular state. In the words of Bertrand Russell[18], rationality is the "just adaptation of means to ends," an approach to be referred to as *formal rationality*. Given that the EPA (or another group) believes that a concentration of X yields a chance of fatal cancer Y, that the EPA desires to produce Y, and that lowering the concentration to X is possible, the rational action is to control concentrations at X. It is rational because the means were consistent with both the stated beliefs and the stated ends. Rationality, then, is located in the entire regulatory process

and can only be judged by someone familiar with that entire process. This does not imply that treating two aspects such as risk assessment and risk management independently is non-rational, only that the rationality of the assessment begins once objectives have been assigned, e.g., the task of risk assessment is to address specific health effects, such as cancer.

4.5. Summary Remarks on Rationality for Risk Analysis

In summary, rationality lies in the quality of the predictions generated by a risk analysis, but also in the means and ends which the risk analysis will serve. The rationality lies in the three separate parts and in the relationship between those parts. The latter view implies that a risk analysis must be structured to yield predictions related closely to existing means for reaching a clear set of ends.

Since logical rigor is associated most closely with questions of belief, practical action with the development of means, and moral reasoning with the choice of ends, people will tend to locate rationality differently depending upon which of these three activities they value most and understand best. Increasing specialization tends to push people towards this intrinsic rationality, since they may be unfamiliar with the relationship between the parts of action. Scientists, for example, may tend to be more familiar with the state of scientific beliefs than with means or ends acceptable to society. Technologists are apt to be more familiar with means than with scientific belief or ends. Public policy practitioners may be familiar with the relationship between beliefs, means, and ends, but not intimately knowledgeable about specific details concerning the parts. They will, as a result, tend towards formal rationality.

The most complete vision of rationality incorporates beliefs, means, and ends, judged both intrinsically and through their relationship. The "complete" rational person chooses rational ends, selects means appropriate to those ends, and then justifies those means and ends

(to the degree possible) by reasoned beliefs. The individual component parts and their relationship are the subject of scrutiny and clear reasoning. Weakness within either a component or in the relationship between components weakens the rationality of a decision and associated action. Rationality is a kind of web where both the individual "nodes" (belief, means, and ends) and the lines stretching between them (the relationships) must be constantly subjected to reason, and, thereby, strengthened. Individuals or groups may be selected to work on specific aspects of the web, but their work must be guided through a more complete conception of rationality. The present research is intended to provide guidelines for understanding of the quality of predictions generated by a risk analysis, and how that quality is related rationally to means, ends, and beliefs.

In summary, intrinsic rationality refers to addressing risk analysis from a singular perspective, such as the rationality of beliefs, the rationality of means, or the rationality of ends, which may align most closely with the tasks and perspectives of the risk analyst, the risk mitigator, and the risk policy maker, respectively. Although each rationality described is important and of interest in itself, they are insufficient to the decision making process as a whole, which must include the relationship between beliefs, means, and ends.

Having described basic features of rationality in a very general sense, it is necessary now to provide more detail. Specifically, we ask *how* a particular belief, means, or end, or the relationship between them, can be supported by reason. Since most of the literature on rationality has been built around beliefs, and since beliefs are the product of a risk analysis, this aspect of rationality will be used as a primary example in the following section. The general ideas, however, apply to all aspects of rationality except as noted.

5. CHARACTERISTICS OF REASONS AND RATIONALITY

This section of the report explores briefly the history of attempts to show why, and how, a given belief or prediction may be claimed to justify adoption of a particular means to a stated end. The various attempts will differ in the manner and degree to which beliefs, means, and ends impact on each other. These differences lead to competing views on the extent to which a risk analysis can be driven entirely by the internal goals of the scientific research community.

Sections 5.1 through 5.9 describe, from a historical perspective, philosophies on the evidential support of belief. Section 5.10 moves to the question of whether belief is adequate. Finally, some recent studies in the history and sociology of science are discussed that help to set the stage for current thought on this topic. The concept of coherence as a goal for the rational human is discussed. A summary section provides a brief listing of the various approaches to the rationality of evidence and belief.

5.1. The Classical Theroy

Discussions of the history of rationality begin with Plato in his depiction of Socrates.[19] This Platonic view has come to be known as the Classical Theory of Rationality, and much of modern philosophy has arisen either in reaction to, or in support of, the classical theory. Most practicing risk analysts presume the theory in supporting the use of risk analysis.

Plato conceived of rationality as the pursuit of Truth. When a person grasps the Truth, he or she also has grasped Beauty, which is the highest end in life. To be rational, then, means to have possession of the Truth. Moreover, it is necessary to realize that the Truth has been located.[20] In this sense, the Classical Theory of Rationality is similar to the rationality of belief discussed earlier. Still, it involves the rationality of ends (since the highest end is truthful belief) and the rationality of means (since it insists on logic as the proper route to truth).

The classical theory has been summarized by Brown[21] and by Agassi[22]. The theory involves five primary attributes by which a belief may be judged rational.

(1) The belief must contain terms which are clearly defined an objective features of the world. There is no room for ambiguity, since the world is not ambiguous.

(2) The belief must be formed on the basis of clearly defined rules of reason, which are taken to be those of logic and mathematics. In short, the belief must be deduced from what already is known by well established rules of reason.

(3) The rules must be universal and applied consistently. As stated by Kant[23], "Act only on that maxim through which you can at the same time will that it should become a universal law." It must be very clear when the rules are to be applied. This attribute is intended to prevent people from intentionally shifting allegiances to beliefs when they are inconvenienced by them.

(4) A belief must be necessary in the logical sense. It must both rest on a firm foundation of established truths and follow deductively from those truths. All persons must come to the same conclusions if they are open to reason.

(5) There must be an algorithm which allows the rules to be applied in a finite number of steps. An example of this is the use of syllogisms or of mathematical formulae. Rules alone are not enough if they cannot be completed in a finite time.

The net result of the classical theory of rationality is to remove human judgment from beliefs. The human acts only as the vessel of belief. He or she has no choice in what to believe, since the belief is dictated by firmly established truths and the laws of logic. There will, therefore, be a complete consensus of belief between all rational people. In a very real sense,

people are not responsible for their rational beliefs and have a guarantee that the beliefs will lead to correct interactions with the world.

The most well developed conception of classical rationality is contained in set theory and its application to a "covering law" hypothesis of science. The object to be analyzed (such as human health and radon) is broken into distinct analytical categories. These categories contain the entities of the world (lungs, radiation, radon, cells, cancer, etc., in our context) and constitute a well defined and finite set. Each entity is said to possess a set of characteristics or attributes (mass, energy, transformation, etc.) which explain the behavior of the entities. Explanations of a particular phenomenon, such as radon yielding lung cancer, then are based on these attributes, in all cases "covered by" the laws of the attributes. For example, it might be stated that all large populations of people receiving X amount of radiation show a fraction, Y, which die of lung cancer. This statement is a "covering law" which covers all phenomena satisfying the condition of being a population receiving X amount of radiation. The rational person faced with radon in the home is concerned with whether his case falls into the set of entities satisfying the above condition. It either does or does not, i.e., each phenomenon is either in the set or out, and the behavior of the phenomenon is certain if it falls into the set. The task of the rational person is to be very clear as to how an entity qualifies to be in the set and to apply the appropriate rules of logic in analyzing the properties and interactions between sets.

This classical view leads to a situation in which rationality and logic are essentially identical. To Plato the world consists of universal, necessary, and eternal truths, so rationality must also possess these attributes. Anything less may be a useful tool for action (Aristotle's practical reasoning) but cannot count as being rational. These are very strong requirements, and very little in life satisfies this classical model of rationality.

5.2. The Skeptical Attack on the Classical Theory

Attacks on the classical view have tended to take the form of skepticism. Even in Plato's time, there was a philosophical school, the Skeptics, who claimed that nothing satisfies Plato's ideal. This led the Skeptics to the claim that no belief could be rational. It is important to note here that this extreme pessimism followed from the Skeptics adherence to the classical view of rationality.[24]

5.3. Empiricist Rationality

Locke[25] hoped to bring rationality back into respect by focusing on experience as providing a base for firm belief. Observations, particularly those from well controlled experiments, would yield beliefs which were a necessary consequence of the observations. If this were the case, the goal of Plato would be reached. Hume[26], however, followed with a skeptical response. As suggested by March[27], rationality requires a decision as to the future actions of the world. It is in the future, after all, that decisions will be carried out. For Hume, no prediction of the future follows necessarily from past observations. Without necessity, skepticism was the proper attitude and rationality was impossible. While Hume agreed with Locke that empirical evidence, i.e., experience, was the best route to knowledge, he would not admit that necessary beliefs ever could be found. Hence, Hume denied claims to rationality.

5.4. Rationalist Rationality

Descartes[28] found Hume's reliance on experience untrustworthy. He rejected the empiricism of Locke and Hume to rely on human insights instead (a bit like Plato). He believed in the capacity of reason to discover truths which were "clearly and distinctly perceived," much as in logic or mathematics. Mathematics was, in fact, his model for rationality. The

rational person discovered, through introspection, a set of axioms about the world. The rules of mathematical logic then were applied to these axioms to yield necessary truths. Since these rules are clearly defined, universal, and led to necessary truths, a rationalist in the Cartesian sense satisfied the requirements of the classical view of rationality. All that was required was the rather questionable belief that the axioms of the world could be obtained through a special form of introspection.

5.5. Logical Positivism

At the beginning of the 20th century, logical positivism[29] was introduced through a philosophical circle in Vienna, known today as the Vienna Circle. Logical positivists returned to Locke in believing that it was possible to begin reasoning with a set of "observation sentences" concerning the objective properties of the world. These sentences would refer entirely to sense impressions, but primarily sight, which was taken to be the most reliable sense. Experimentation would provide those impressions. Being purely statements of observation, a rational person could be "positive" about those beliefs. The rules of logic then could be applied to deduce any new predictions about the world (hence, the name "logical positivism"). The only trick was to produce observation statements which were necessarily true statements about the objective features of the world. To be rational meant to believe a statement only if it was an observation statement or followed deductively from observation statements. As such, logical positivism is a form of the classical view of rationality, wedded to a strong idea of empiricism.

5.6. Rational Skepticism

Karl Popper was initially attracted to the Vienna Circle, but broke from that line of reasoning. His primary complaint was that no statement about the future could follow

necessarily from past observations (the contribution of Hume). Such statements were, instead, mere hypotheses requiring further investigation.[30] The proper attitude towards all statements, including those of science, was skepticism. Still, Popper rejected the classical skeptic's claim that people could not be rational. He argued that rationality was to be found in the *process* by which hypotheses were subjected to tests. Central to his idea of rationality is criticism of an idea, which included attempts to show that the idea is wrong. A rational person attempts to find evidence which *falsifies* beliefs. As in logical positivism, this evidence was taken to be observational experience. The rational person adopted the belief which had survived the most stringent attempts to falsify it. Competing beliefs could not be shown to be true, but it was possible to talk of the "truth content" of beliefs by referring to their success in avoiding refutation or falsification. Rationality, in Popper's view, is a process of criticism in the form of attempts at falsification and acceptance of the belief most resistant to criticism.

Another key point on which Popper and the logical positivists disagreed is the role of "confirming" evidence in choosing beliefs. As mentioned previously, the positivists assumed that observation of a phenomenon, such as a risk of Y, counted as a positive reason for believing any theories which predicted Y. The observation in a sense confirmed these theories. A rational person chose the belief which had been the most highly confirmed or "verified." This test of verification was made against the full range of existing observations deemed to be relevant to a belief. Popper countered that it was not enough for a belief to be confirmed by an observation, since it might be the case that *any* belief would be consistent with that observation. It might also be the case that a belief was so vague that *any* observation could be construed to confirm it (this Popperian argument is often made against multistage theories of carcinogenesis). This led him to his rationality of falsification, in which there must be a strong chance that a theory would fail to be confirmed by an experimental test. Popperian epidemiologists have, for example,

argued that the multistage theory of carcinogenesis is so flexible that it could fit any set of data. There is no way, therefore, to say that the theory has the potential to be falsified, and it should be rejected as unscientific.

5.7. Probabilistic Rationality

Regardless of whether verification or falsification is adopted as a rational process, there is the problem of ambiguous, variable and conflicting evidence. All beliefs will have some evidence in their favor and some against. In addition, observation will be limited to finite samples which, as Popper correctly points out, may miss important features of reality. Recognition of this situation gives rise to the probabilistic theory of rationality. Here, evidence does not necessarily verify or falsify a given belief. It lends, instead, a "degree of belief" based on some idea of the probability that a piece of evidence would have been produced if the belief were correct.

This degree of belief has taken two main forms in the literature on rationality. The first is classical statistics.[31] Statistical properties of the method of observation are used to estimate the probability that an observed result would be obtained if a belief were true. This probability is taken to be an objective property of the method of observation, so the resulting assignment of probability satisfies the classical requirements for objectivity, necessity, and universality. A rational person chooses the belief which is assigned the largest probability based on the existing observational evidence. All persons who agree with the statistical assumptions (which themselves may be open to question) will assign the same degree of confirmation or belief.

Contrasting with this objective approach is a subjective approach to probability often referred to as Bayesian confidence (although it is possible to establish a quasi-objective Bayesian methodology[32]). Bayesians recognize that new beliefs follow from a mixture of prior beliefs and

new observational data. In a methodology to be discussed in a later report, the confidence in a belief is adjusted historically as new data are presented to the mind. This confidence is said to be a measure of the state of mind of a person, rather than an objective property of the observed phenomenon or the method of observation. Still, the Bayesian approach can be made to yield consistent and necessary estimates of confidence based on a formal set of axioms[32], thereby satisfying some of the requirements of the classical view of rationality. The rational person then adopts the belief associated with the highest confidence.

5.8. The Problem of Incommensurability

Common to all of the views discussed above is an assumption that beliefs, such as scientific theories, can be compared simultaneously to a common set of observations. Once the observations are specified, the degree of verification or falsification can be assigned and the beliefs ranked accordingly. Kuhn[33], however, denied this claim. Using scientific theories as the main example of a potentially rational activity, he asserted that theories are basically incommensurable. In other words, the theories themselves specified what to look for in the world and how those observations would count as evidence. People working under different theories observed the world in different ways and spoke differently about the world. As a result, there was no fixed set of observations against which competing beliefs could be assigned measures of confirmation or falsification. Beliefs must be seen instead as systems of belief, each of which set up traditions of research but basically could not be compared. It was possible to be rational *within* a scientific research tradition, but all such traditions were internally rational so long as they remained open to change in light of the evidence they uncovered.

Kuhn's position on the rationality of science was misused by both his opponents and supporters, prompting him to reply in a second book.[34] The primary mistake made by both

groups was the assertion that his position leads to complete relativism. If all scientific theories choose the data which must be explained, and how those data are to be explained, why not say that all areas of human knowledge are equally valid? After all, all subject areas, e.g., science, religion, art, specify the experiences to be confronted, the methods of exploration, and the rules of explanation. Since these features only make sense *within* the language of the theories, there is no way to step outside of them all and compare them on a common basis. Science, mythology, religion, and art simply become competing systems of thought. Each is internally consistent and none may be rejected through appeal to higher criteria. The same may be said of competing scientific theories.

This view found favor both in the social sciences and in the philosophy of Paul Feyerabend[35]. Feyerabend argued that scientific theories were completely incommensurable and internally coherent. There was not, therefore, any way to reject one in favor of any other. Scientific choice between theories was profoundly irrational and the only reason scientists chose one over the other was to be found in the exercise of power, prestige, etc. To be rational, a society must give up the quest for selection between scientific theories. All theories, whether scientific or not, must be encouraged to prosper on an equal footing. When the time came for decision, democracy must be the method for choice. Kuhn's critics saw in Feyerabend the natural culmination of the thesis of incommensurability, leading inexorably to irrationality, or non-rationality, and ultimately to an anarchy of beliefs.

This turn of events was troubling to Kuhn. Within social science and anthropology, the Kuhnian idea of a paradigm came to mean any model of the world. Since most ways of thinking counted as a model of the world, most ways of thinking satisfied the (false) picture of a paradigm and, therefore, constituted a perfectly acceptable theory of the world. But this misuse of the idea of a paradigm greatly simplified and distorted Kuhn's original intent, although part

of the blame lies in the vagueness of his original writings. Paradigms were intended to provide a concrete means of confronting scientists with empirical tests of their theories. Different theories might specify different tests to be performed, but all proper paradigms gave rise to experiments which held the chance of verifying or falsifying the theory through tests against experience. Not every theory of the world satisfied this criterion, so not every theory of the world was rational. In addition, researchers within a theory were capable of admitting that their theory was not working very well by their own rules. There even were many observations which all theories agreed were important tests of any theory. Subjectivity and relativism were not natural consequences of his view on rationality. Still, it is important to bear in mind his important observation that researchers working within different theories may look for different bodies of evidence in supporting those theories.

5.9. Tests of Theories Other Than Empirical Tests

Kuhn shared a common ground with Aristotle, Locke, Popper and the logical positivists in adopting empiricism as the test of a belief. The proof of a theory was in the degree to which its assumptions could be observed or in its ability to correctly predict the results of an observation. Laudan[36] agreed with this claim but added another criterion for rational confidence. At times, a theory solves *conceptual* problems, which are quite different from empirical problems. Consider, for example, a theory of radiation carcinogenesis which proposes that radiation damages DNA and turns on an oncogene. This oncogene then produces cancer. The theory provides an explanation which includes the role of oncogenes. If scientists had been wondering how oncogenes fit into the picture of radiation and cancer, the new theory may be said to solve a conceptual problem. This solution counts in favor of the theory even if it does not yet lead to increased precision in predicting the incidence of cancer. With Laudan,

therefore, rationality takes on a new measure. A rational theory explains, in a conceptual sense, the most significant problems remaining in a field. While the ability of a prediction to "fit" data is an empirical problem dictated by the rules of observation and statistical reason, the conceptual ability of a theory is judged by humans. It is the human who assigns relevance and importance to a conceptual problem, requiring an act of judgment. It is the human who decides how well a theory explains a particular problem or puzzle.

A similar tact was taken by Longino[37]. She admitted that rules of reason might well apply once a set of data were chosen as a test of a theory. Choosing the data which were relevant as a test was, however, a process buried in human judgment. The cell biologist studying the transformation of cells in vitro following irradiation might assume the results were highly relevant to predictions of cancer in humans. Any theory of cancer would, therefore, be required to explain and predict the in vitro experiments. Epidemiologists might argue the opposite, asserting that the in vitro studies were irrelevant to theories of cancer. There is little hope of a logical resolution to the decision, since assigning relevance to an observation presupposes the theory which gives relevance to begin with. In this sense, Longino's insights are similar to those of Kuhn. Far from being an abstract exercise in philosophy, the concerns of Longino were at the heart of a debate on the effects of radiation on human cancer.[38]

5.10. The Rationality of Crafting in Science

All of the philosophies discussed to this point have held a common feature. They were concerned with the degree to which evidence might support a belief. The belief was precise and the only question was whether it was true or false. Laudan[36] moved the focus slightly and asked whether a belief is *adequate*. A key feature of modern philosophy of science is increased use of the concept of approximation. No theory or belief is said to be completely true or false. Parts

of the world are missing from models. The mathematical relationships between the parts are understood partially. Even measurements of agreed upon relationships are imprecise. If truth and rationality require complete empirical success for a theory, then no theory is true and no one is rational.

And yet, scientific theories can be said to form a rational basis for belief. All that is required is a change of emphasis on the need for complete predictive success. A successful theory then predicts adequately within the guidelines set up by a community. Human judgment is needed to define the term "adequate." This judgment cannot be avoided and usually will be related to the ends to which the theory will be used. Still, once the idea of adequacy is given clear expression, an imperfect theory may be said to provide a rational basis for belief.

A similar idea was advanced by Ravetz[39], who likened science to a craft. Adequacy could be thought of as being analogous to "tolerance" in engineering. If two parts must fit together within an engine, there is no need for them to be perfect fits. They must, instead, fit together within a certain margin of error (called the tolerance). Greater accuracy (beyond the desired tolerance) does nothing to satisfy the goal of making a working engine. Lower accuracy may, however, result in the engine shaking itself apart. As with Laudan, the rationality of belief for Ravetz requires an idea of how a theory is to be used and how adequate a theory is to that use. Ends, means, and beliefs converge, therefore, in the act of crafting.

This idea of crafting is carried even further by Polanyi[40], Heidegger[41] and Rouse[42]. Their argument is that science is, first and foremost, a way of *doing something*. As mentioned earlier, the proof of science (or of a scientific belief) might be said to lie in "knowing how," in practice, rather than "knowing that." A proper belief is simply a conceptual representation of a method for doing something. Experiments test the ability to produce a particular physical state in the world, not the truth of beliefs. Beliefs (such as theories) are ways humans have of talking about

successful actions. That, and only that, makes a belief rational or irrational. The rationality of belief and of means shade over one into the other.

The focus on crafting and doing tends to negate some of the original concerns of Popper. Popper discounted confirming evidence both because a theory or model might agree with reality for the wrong reason and because they might be adjustable to fit almost any set of experiences (such as experimental data). This is, indeed, a major problem if it is required that a theory or model be a truthful explanation of the world. If, however, their role is to conform only to the surface features of the world (i.e., to mimic the world), then truthful explanations are not so important and confirmation implies that the theory or model at least *functions* like the world (even if for the wrong reasons). This functional ability may be thought to make the theory or model a perfectly rational basis for *action*, if not *belief*.

5.11. Scoiological Views of Science

This brings to a fairly complete end the discussion of rationality and belief. Two topics remain to be addressed, however, if only briefly. The first concerns recent studies in the history and sociology of science.[42,43] The purpose of philosophy of science has been to explain why it was rational for scientists to adopt a particular belief in light of evidence existing at some moment. This might require what is known as a "rational reconstruction" of the actual history of a science, but it still would be possible to show that a belief might have been rational according to the criteria discussed in earlier paragraphs. Someone using the results of a risk analysis typically is more concerned with whether a belief can be made rational, not whether a specific analyst was in fact rational in developing that belief. Sociologists of science, however, denied that any such rational reconstruction was either possible or relevant to the history of science (or of risk analysis).

These sociologists directed the focus away from the relationship between evidence and beliefs and onto the social and physical setting of science. Acceptance or rejection of beliefs was said to remain loosely based in evidence and mathematical reason, but the main driving forces were prestige, social relationships, funding, strong personalities, etc. These factors influenced the direction of research, the methods chosen for research, the assignment of relevance to evidence, and so on. Since real risk analyses come out of actual historical circumstances, rather than rational reconstructions of history, it seems wise to pay heed to the lessons from sociology and history of science, if only to work harder at strengthening the influence of more classical notions of rationality. The lesson from sociology is that risk estimates from "expert committees" may be driven more by the dynamics of the committee than by the evidence itself.

5.12. Foundations Versus Coherence

Finally, there is an issue related to the proper metaphor to be used in picturing rational belief. Philosophers in the classical vein tend to follow the idea of foundationalism. They conceive of knowledge and belief as resting on a set of foundations or well established "facts' of the world. All other beliefs are built on these foundations. Quine[44] challenged this view of belief. He noted that no philosophy had yet discovered a firm foundation beyond dispute. In the place of foundations he inserted the idea of a "web of belief." Since no belief acted as a foundation, all beliefs must be mutually supporting. Changes in one belief then would affect all of the others. A rational person strives for *coherence* between beliefs. Conflicting beliefs are rooted out and reconciled in a constant process of revision. No single belief is immune to this revision, even (as Maxwell also asserts[14]) metaphysical beliefs about the ends and methods of science.

5.13. Summary

A very brief listing of the various approaches to the rationality of evidence and belief is given below.

1. The Classical Theory--a risk analysis is rational if the predictions are necessarily true. Truth is the only goal, regardless of whether that truth helps to meet other goals (it is *presumed* that it does). This necessary truth may be founded on empiricism or rationalism.

   A. Empiricism--necessary truths in risk analysis are obtained by restricting the predictions to those which have been observed. A variant of this approach is logical empiricism or logical positivism.

   B. Logical Positivism--necessary truths in risk analysis are obtained by restricting the predictions to those that are *deduced* (logically) from observations.

   C. Rationalism--necessary truths in risk analysis are obtained by restricting the predictions to those that are *deduced* (logically) from "clearly and distinctly perceived" (i.e., intuited) insights, as in the axiomatic approach to mathematics.

2. The Skeptical Approach--no belief is a necessary truth. Therefore, rationality is not possible.

3. Critical Rationality--necessary truths cannot be attained from a risk analysis. The proper attitude is one of skepticism. Still, rationality is possible by requiring that beliefs constantly be questioned and debated. This debate occurs through the testing of beliefs against evidence. The test may be one of falsification or verification. The evidence usually is taken to be observational, and the test to be correspondent to the observation. At any moment, it is possible to compare all beliefs against a single body of evidence.

4. Paradigmatic Rationality--necessary truths cannot be attained from a risk analysis. The judgments of belief (as in #3) must be made on the basis of evidence, but the evidence differs between different beliefs (i.e., between scientific theories). Rationality implies critical discussion, but the discussions take place *within* research traditions, with these discussions being incommensurable.

5. Conceptual Rationality--theories used in risk analyses are not judged purely on correspondence to evidence, but on the degree to which the theory resolves *conceptual* difficulties in a field. Empirical tests (as in #3 and #1B) are important, but so is conceptual clarity and coherence.

6. Rationality of Evidential Relevance--rationality in risk analysis requires that evidence (regardless of whether it is empirical, conceptual, or the result of insight) be judged *relevant* to a problem. This judgment is not a necessary truth and must be made by researchers using human skills. The judgment of relevance becomes part of the background assumptions (often hidden) adopted by a particular analyst. Complete rationality requires a critical discussion of these judgments.

7. Sociological Rationality--the goal (usually hidden) of the risk analyst is to maximize power, prestige, etc. Rationality requires that these goals be recognized and a social system adopted to modify their effect on the analysis.

8. Instrumental Rationality--a risk analysis is rational only if it produces explicit means for mitigating the risk. Strong beliefs are not sufficient if they do not lead to demonstrable solutions.

6. CONCLUDING REMARKS

The preceding chapters provide an overview of the nature of rationality and indicate properties required of a framework for it to reflect this nature in risk analyses of environmental carcinogens. Specifically, the framework (to be developed in the next part of this report) must have the following properties:

(1) It must be capable of depicting how inferences of risk are related to premises concerning probability, evidence and the physical world (satisfying the deductive ideal of rational analysis).

(2) It must incorporate the full range of evidence typically brought to bear on a risk analysis for carcinogens (satisfying the ideal of rational coherence).

(3) It must be capable of making explicit the relevance of particular bodies of evidence to inferences of risk (satisfying the requirement of rational analysis of evidential relevance).

(4) It must be capable of allowing an explicit incorporation of potential links between beliefs, means, and ends (satisfying the requirement of formal rationality). These links will depict how specific effects are selected for analysis contingent on specified ends and how specific qualities of evidence are weighted according to the degree to which means (or methods of mitigation) are considered.

(5) It must be capable of displaying the critical points at which human judgments must be made of evidential strength and relevance, and of relating these

judgments to specific assumptions concerning the nature of rational belief (thereby aiding researchers in assessing their state of confidence).

(6) It must be capable of demonstrating how competing risk analyses arise from differing explicit assumptions (aiding in the explication and resolution of disputes).

(7) It must be susceptible to use as a tool for producing either qualitative or quantitative judgments concerning the strength of inferences of risk (thereby avoiding necessary commitment to specific schools of thinking about evidential strength, the nature of probability, etc.).

(8) It must provide an explicit link between judgments made by researchers in a wide range of scientific research, as well as judgments made by parties concerned with the less specific issues of evidential reasoning, appropriate ends, and so forth (satisfying the goal of game rationality).

The next part of this report develops a framework for discourse on hazard identification with the eight properties described above. Additional useful concepts, such as a taxonomy of claims of carcinogenicity, are also described and implemented in the framework. The result is a seven-step guide to aid the individual decision maker in drawing inferences of carcinogenicity and to facilitate discourse and aid conflict resolution between persons of differing opinions.

Table 0 has been constructed to aid the reader in making connections between this part of the report and the next, and to make explicit where and how the concepts and principles of rationality discussed in the preceding chapters are implemented in the seven-step guide for applications. The table is located below to facilitate easy reference while reading Part II, which follows.

TABLE 0. A SUMMARY OF IDEAS IN CHAPTER 1 AND THEIR RELATIONSHIP TO THE FRAMEWORK OF ANALYSIS PROVIDED IN CHAPTER 2.

Philosophical Principle	Summary Description	Relevant Pages in Chapter 1	Role within the Framework of Analysis in Chapter 2
Conceptual Rationality	An attempt to provide consistent and clear meanings to terms in an analysis.	12, 15, 39	Useful in establishing meaning of "claims" in Working Tables 3, 6, 7; as well as all other terms in the analysis.
Logical Rationality	An attempt to apply rules of deductive reason in arriving at the results o f an analysis equivalent to classical rationality.	12, 22, 38-41, 42	Useful in determining the assignment of epistemic value to the relevance strategies in Working Tables 3, 6, 7; and to the column and overall summary entries in these tables.
Methodological Rationality	The habit of questioning beliefs used in an analysis. Also termed process rationality.	12, 22, 43	Useful in determining the assignment of epistemic value to the relevance strategies in Working Tables 3, 6, 7.
Epistemiological Rationality	An exploration of the evidential support for beliefs, and of the principles of by which the support is judged.	12, 14	Useful in determining the assignment of epistemic value to the relevance strategies in Working Tables 3, 6, 7.
Ontological Rationality	The insistence that only entities demonstrated convincingly to exist be employed in the reasoning of an analysis.	12, 39	Useful in determining the data items to be incorporated into Working Table 4.
Valuational Rationality	The insistence that the process of analysis reflect the most valued aspects of human intellectual activity and of the ends of the analysis.	12, 16, 23	Useful in determining the degree of intellectual obligation assigned to relevance strategies in Working Tables 3, 6, 7; as required by column summary.
Practical Rationality	The insistence that an analysis be a practical means to action.	12	Useful in determining whether the different "claims of carcinogenicity" categories are desired goals of the analysis (Working Tables 3, 6, 7).
Evidential Relevance	The ability of evidence, in the context of background assumption to have bearing on a given belief.	14	Useful in establishing Context Type (Working Table 1 and Step 2 of Fig. 4); Causality in Working Table 2; and relevance strategies (Working Tables 3, 6, 7).
Empirical Success	The ability of a theory to predict the observable properties of a phenomenon. See also logical positivism (p. 42).	18, 41, 42	Useful in determining the assignment of epistemic value to the direct empirical, semi-empirical, and theory-based-inference relevance strategies in Working Tables 3, 6, 7.

(continued on following page)

Table 0. (continued)

Philosophical Principle	Summary Description	Relevant Pages in Chapter 1	Role within the Framework of Analysis in Chapter 2
Conceptual Success	The ability of a theory to provide a causal explanation of observable properties of a phenomenon.	18	Useful in determining the assignment of epistemic value to the theory-based inference relevance strategy in Working Tables 3, 6, 7. Also useful in assigning causality in Working Table 2.
Hard look	The claim by an analyst that the assembled evidence constitutes a reasonably complete body reflecting the full body of evidence that would be available given no restrictions of time or money.	19	Determination of the value of data completeness in Working Tables 2, 4.
Correspondence	The philosophical principle that a theory is established to be true by its ability to predict a given observation.	18, 51-52	Useful in determining the assignment of epistemic value to the theory-based-inference relevance strategy in Working Tables 3, 6, 7.
Coherence	The philosophical principle that a theory or belief is established to be true if it is not inconsistent with other beliefs held by the analyst.	20, 51-52	Useful in determining the assignment of epistemic value to the theory-based inference relevance strategy in Working Tables 3, 6, 7. Also useful in determining the assignment of epistemic value for column and overall summaries in these tables. Also useful in ensuring that value of intellectual obligation is consistent across analyses.
Fides	Those beliefs taken by the analyst to require no attempt at evidential justification.	20, 22	Useful in determining the strength of background assumptions appearing in relevance strategies of Working Tables 3, 6, 7.
Epistemic Status	The weight of the evidence brought to bear in supporting a belief.	21	Identical to assignment of epistemic value to relevance strategies in Working Tables 3, 6, 7.
Adaptive Rationality	The belief that a rational system of analysis should contain flexibility to accommodate new insights/knowledge. See also posterior rationality (p. 24).	23, 24	It is useful to reflect on this concept in determining the degree to which categories of carcinogenicity claims will be tied to specific theories of carcinogenesis.

(continued on following page)

Table 0. (continued)

Philosophical Principle	Summary Description	Relevant Pages in Chapter 1	Role within the Framework of Analysis in Chapter 2
Selective Rationality	The set of principles by which individuals are selected for particular tasks and/or judgments.	23	Useful in determining the assignment of epistemic value to the existential insight relevance strategy of Working Tables 3, 6, 7.
Rationality of Belief	The ability to demonstrate that beliefs are justified by available evidence.	29-35	Useful in reflecting on the epistemic status of claims in Working Tables 3, 6, 7.
Rationality of Means	The demonstration, through reference to beliefs, that a given means is likely to produce the desired ends.	29-35	Useful in reflecting on whether a chosen category of claims of carcinogenicity will be useful in reaching the ends of the analyst and/or policy-maker.
Instrumentalism	The philosophical position that theories are established to be true if they allow practical action.	33	Useful in determining the assignment of epistemic status to the theory-based-inference relevance strategy of Working Tables 3, 6. Also useful in assigning degree of intellectual obligation to the relevance strategies in these tables.
Realism	The philosophical position that theories are established as true if they contain reference only to entities and relationships that have ontological status (see ontological rationality).	32	Useful in determining the assignment of epistemic status to the theory-based inference relevance strategy of Working Tables 3, 6. Also useful in determining allowed data items in Table 2.
Rationalism	The philosophical position that beliefs are established as true if they are clearly and distinctly perceived by the human mind, or are obtained from rationally justified beliefs using the rules of deductive logic.	42	Useful in determining the assignment of epistemic status to the various relevance strategies in Working Tables 3, 6, 7.
Contextual Rationality	The philosophical principle that focuses on a particular aspect of an analysis should not interfere with the overall goal of the analysis.	23	Useful in determining when examination of the evidence in a data category is to be considered reasonably complete (Working Tables 2, 4).

REFERENCES (PART I)

1. B. Russell, "Philosophy and Politics," in *Unpopular Essays*, Simon and Schuster, NY, 1950.

2. K. Popper, *The Open Society and Its Enemies*, Routledge and Kegan Paul, 1951.

3. This point is raised by W. Berkson, "Skeptical Rationalism" in *Rationality: The Critical View*, ed. by J. Agassi and I. Jarvie, Martinus Nijhoff Publishers, Dordrecht, p. 27, 1987.

4. M. Cohen, *Reason and Nature*, Free Press, Glencoe, Illinois, 1953.

5. W. Newton-Smith, *The Rationality of Science*, Routledge and Kegan Paul, London, 1981.

6. J. Rouse, *Knowledge and Power*, Cornell University Press, Ithaca, NY, 1987.

7. W. Mills and D. Egan, Jr., "Risk Assessment and Policy," in *Environmental Radon*, ed. by C. Cothern and J. Smith, Jr., Plenum Press, NY, pp. 273-295, 1987.

8. M. Friedman, *Rational Behavior*, University of South Carolina Press, Columbia, SC, 1975.

9. K. Irani, "Introduction: Modes of Rationality," in *Rationality in Thought and Action*, ed. by M. Tamry and K. Irani, Greenwood Press, NY, pp. xi-xx, 1986.

10. R. Wollheim, "Wish-fulfillment," in *Rational Action*, ed. by R. Harrison, Cambridge University Press, Cambridge, pp. 47-60, 1979.

11. B. Fischoff, P. Slovic, S. Lichtenstein, S. Read and B. Combs, "How Safe is Safe Enough? A Psychometric Study of Attitudes Towards Technological Risks and Benefits," *Policy Sciences*, 8, 127-152, 1978.

12. J. Cook and D. Egan, Jr., "Mitigation," in *Environmental Radon*, ed. by C. Cothern and J. Smith, Jr., Plenum Press, NY, pp. 249-272, 1987.

13. D. Hume, *A Treatise of Human Nature*, ed. by L. Selby-Bigge, Clarendon Press, Oxford, 1888.

14. N. Maxwell, *From Knowledge to Wisdom*, Basil Blackwell, Oxford, 1987.

15. D. Hamlyn, "Empiricism," in *The Encyclopedia of Philosophy*, MacMillan Publishing Co., NY, pp. 499-505, 1967.

16. M. Scriven, *Primary Philosophy*, McGraw Hill Book Co., NY, pp. 10-49, 1966.

17. See, for instance, the discussion in A. MacIntyre, *Whose Justice? Which Rationality?*, University of Notre Dame Press, Notre Dame, Indiana, 1988.

18. See the discussion of Bertrand Russell in *Encyclopedia of Philosophy*, McGrawhill, New York, pp. 417-430, 1974.

19. See, for instance, the discussion in S. Nathanson, "The Classical Ideal," in *The Ideal of Rationality*, Humanities Press International, Atlantic Highlands, NJ, pp. 3-14, 1985.

20. This point is brought out, for example, on p. 35 of H. Brown, *Rationality*, Routledge, London, 1988.

21. H. Brown, *Rationality*, Routledge, London, 1988.

22. J. Agassi, "Theories of Rationality" in *Rationality: The Critical View*, ed. by J. Agassi and I. Jarvie, Martinus Nijhoff Publishers, Dordrecht, pp. 249-263, 1987.

23. I. Kant, *Groundwork of the Metaphysic of Morals*, trans. by H. Paton, Harper and Row, NY, 1956.

24. This point is raised in J. Agassi, "Theories of Rationality" in *Rationality: The Critical View*, ed. by J. Agassi and I. Jarvie, Martinus Nijhoff Publishers, Dordrecht, pp. 249-263, 1987.

25. J. Locke, *An Essay Concerning Human Understanding*, reprinted by Dover Publications, Inc., NY, 1959.

26. D. Hume, *An Enquiry Concerning Human Understanding*, ed. by L. Selby-Bigge, Clarendon Press, Oxford, 1902.

27. J. March, "Bounded Rationality, Ambiguity, and the Engineering of Choice," *The Bell Journal of Economics*, pp. 587-608, 1978.

28. G. Anscombe and P. Geach, *Descartes: Philosophical Writings*, Edinburgh, 1954.

29. Discussions on this topic are numerous, and the reader might consult H. Putnam, *Reason, Truth and History*, Cambridge University Press, Cambridge, 1981.

30. K. Popper, *The Logic of Scientific Discovery*, Hutchinson, London, 1959.

31. For basic review and references, see "Scientific Method and Statistics" in *Encyclopedia of Statistical Sciences*, Vol. 8, John Wiley & Sons, NY, 1981.

32. For basic review and references, see "Bayesian Statistics" in *Encyclopedia of Statistical Sciences*, Vol. 2, John Wiley & Sons, NY, 1981.

33. T. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, 1962.

34. T. Kuhn, *The Essential Tension*, University of Chicago Press, Chicago, 1977.

35. P. Feyerabend, *Against Method*, New Left Books, London, 1975.

36. L. Laudan, *Progress and Its Problems*, University of California Press, Berkeley, 1977.

37. H. Longino, *Science as Social Knowledge*, Princeton University Press, NJ, 1990.

38. National Academy of Sciences, "The Effects on Populations of Exposure to Low Levels of Ionizing Radiation: 1980," National Academy Press, Washington, 1980.

39. J. Ravetz, *Scientific Knowledge and Its Social Problems*, Oxford University Press, Oxford, 1971.

40. M. Polanyi, *Personal Knowledge*, University of Chicago Press, Chicago, 1958.

41. M. Heidegger, *Being and Time*, translation by Peter Hertz, Harper and Row, NY, 1962.

42. B. Barnes and D. Bloor, "Relativism, Rationalism and the Sociology of Science," in *Rationality and Relativism*, ed. by M. Hollis and S. Lukes, MIT Press, Cambridge, MA, pp. 21-47, 1982.

43. B. Barnes and D. Edge, eds., *Science in Context*, MIT Press, Cambridge, MA, 1982.

44. W. Quine and J. Ullian, *The Web of Belief*, Random House, NY, 1970.

45. M. Bunge, "Seven Desiderata for Rationality," in *Rationality: The Critical View*, ed. by J. Agassi and I. Jarvie, Martinus Nijhoff Publishers, Dordrecht, pp. 5-16, 1987.

46. L. Wittgenstein, *Tractatus Logico-Philosophicus*, Kegan Paul, Trench, Trubner & Co., NY, 1922.

47. L. Wittgenstein, *Philosophical Investigations*, Basil Blackwell, Oxford, 1953.

48. S. Nathanson, *The Ideal of Rationality*, Humanities Press International, Atlantic Highlands, NJ, 1985.

49. R. Carnap, "Testability and Meaning," in *Readings in the Philosophy of Science*, ed. by H. Feigl and M. Brodbeck, Appleton-Century-Crofts, NY, pp. 47-92, 1953.

50. H. Simon, "Rational Choice and the Structure of the Environment," *Psychological Review*, 63, 129-138, 1956.

51. N. Rescher, *Rationality*, Clarendon Press, Oxford, 1988.

52. W. Berkson, "Skeptical Rationalism," in *Rationality: The Critical View*, ed. by J. Agassi and I. Jarvie, Martinus Nijhoff Publishers, Dordrecht, pp. 21-44, 1987.

53. M. Hesse, *Revolutions and Reconstructions in the Philosophy of Science*, University of Indiana Press, Bloomington, 1980.

54. R. von Mises, *Probability, Statistics and Truth*, Dover Publications, NY, 1957.

55. D. Crawford-Brown and N. Pearce, "Sufficient Proof in the Scientific Justification of Environmental Actions," *Environmental Ethics*, 11, 153-167, 1989.

56. A. Sen, "Informational Analysis of Moral Principles," in *Rational Action*, ed. by R. Harrison, Cambridge University Press, Cambridge, pp. 115-132, 1979.

57. W. James, *The Will to Believe*, Dover Publications, Inc., NY, republished in 1956.

58. S. Toulmin, *The Uses of Argument Analysis*, Cambridge University Press, Cambridge, 1958.

59. Committee on the Institutional Means for Assessment of Risks to Public Health, Commission on Life Sciences, and the National Research Council, *Risk Assessment in the Federal Government: Managing the Process*, National Academy Press, Washington, D.C., 1983.

60. U.S. Environmental Protection Agency, *The Risk Assessment Guidelines of 1986*, Washington, D.C., 1987.

PART TWO

ASSESSING EVIDENTIAL SUPPORT FOR CLAIMS OF CARCINOGENICITY: ESSENTIAL ELEMENTS AND A FRAMEWORK FOR APPLICATION

1. INTRODUCTION

1.1. Background and Objective

Scientific evidence of the last ten years increasingly suggests that carcinogens differ according to their chemical or biological properties and that no single cancer mechanism is universally applicable within all contexts or biological settings. Substances may differ in the mechanism by which they bring about cellular conversions, the number of stages necessary for producing cancer, and the particular conversions associated with a specific substance. For example, it is widely accepted that some cancers arise from nongenotoxic origins (Barrett, 1987; Hecker, 1984; Trosko and Chang, 1988) and that a substance may contribute to cancer development without being a complete carcinogen itself (Barrett and Wiseman, 1987). This suggests that the carcinogenicity of a substance may be a product of both the substance and the conditions (context) under which it acts.

Consequently, the task of hazard identification has expanded as well, from simply addressing whether a chemical is a carcinogen in some sense to a broader assessment of carcinogenic potential, which includes conditions and mechanisms by which it may modify or enhance cancer development either alone or in concert with other substances. The need for a more detailed taxonomy is reflected in the ambiguity of the phrase "potential to increase the incidence of cancer..." cited earlier (U.S. EPA, 1986). The term "potential" may refer to the fact that a claim to carcinogenicity may not have been fully proven by the available evidence, or it may refer to the fact that a substance yields an increase of cancer only under a prescribed set of conditions. These two meanings of the term "potential" must be separated for clarity. The

particular ends reached by more detailed taxonomies are (1) the ability to warrant a claim to carcinogenicity through reference to conceptual understanding (Laudan, 1977), (2) the ability to employ biophysical data other than direct observations of cancer incidence in warranting a judgment of carcinogenicity, (3) the ability to cite the context (species, dosing regimen, etc.) within which a substance will be capable of inducing carcinogenic effects and (4) the ability to address the effect of simultaneous exposures to multiple substances possibly acting by different routes in a non-additive manner (although this end requires considerations well beyond those addressed in the present report).

The published literature related to biological aspects of cancer mechanisms is abundant and much of it has potential implications for hazard identification or other aspects of risk assessment. It is difficult, however, to conceptualize a methodologic framework that extracts and integrates that potential in a rational and reasonably comprehensive manner without imposing undue and contentious theories of carcinogenicity onto the analyst. Cancer research is ongoing. Experimental evidence of varying strength and relevance (more on this term later) to human carcinogenicity gives birth to new hypotheses and conjectures, resulting in turn in the formation of new if incomplete paradigms of cancer mechanisms. But the task of identifying environmental carcinogens cannot wait for perfect understanding. What principles, then, would facilitate this task with maximal use of available information? Toward what objective should those principles be directed? How is this objective related to a taxonomy of claims of carcinogenicity?

It is assumed that the objective of hazard identification is to reach the most informed judgment about the carcinogenic potential of a substance of interest *vis-a-vis* the observational evidence and the current state of knowledge relevant to that judgment. An "informed judgment" is taken to be one characterized by reflection on (1) the complete body of data available on a

given substance, (2) the complete body of conceptual schemes by which the data are brought to bear in justifying any of the possible claims, (3) the "strength" of the data and validity of conceptual schemes and (4) the uncertainties introduced by the existence of contradictory data and/or conceptual schemes. As discussed earlier, "carcinogenic potential" is taken to include consideration of both the epistemic status of a claim to carcinogenicity (i.e. how well such a claim may be supported) and the context within which a substance yields an increase in cancer. This suggests several principles for implementation of a process leading to informed judgments on claims to carcinogenicity: complete assembly of observational evidence, assessment of each source of evidence for its quality (reliability) and relevance (to the task of justifying one or more of the possible claims of carcinogenicity), and evaluation of the coherence of conclusions from the available lines of inference.

The concepts of observational evidence and of its quality are well established in the existing scientific literature and are not discussed at length here (more in Chapter 3). It simply is noted that risk analysts must first establish an observational base (called the "set of observation statements" in philosophy of science, but referred to as the "observations" or "data" here) which the analysts determine to be sufficiently reliable to justify the observations' use in subsequent lines of reasoning leading to claims of carcinogenicity. Clearly, no conclusion of carcinogenicity can be stronger than the base of empirical data on which that conclusion must ultimately rest unless non-empirical epistemic foundations are deemed an appropriate scientific concept. The strength of this observational base increases as (1) the variety of relevant empirical data increases, (2) the quality of each specific body of empirical data increases and (3) the existing empirical data, taken as a whole, display coherence rather than contradictory observations.

The concept of relevance is less well established in the scientific literature, although it is an important concept in logic and philosophy of science (Longino, 1990). In most cases, observations available to the analyst are not direct observations of cancer in the human population under conditions of interest (i.e. in the desired context). They might, instead, be observations of DNA adducts, hyperplasia, increased cancer incidence in other species, etc. The role of such data in supporting claims to carcinogenicity depends critically on the introduction of premises into a line of reasoning leading from the data to the conclusion of carcinogenicity. These premises constitute a set of background assumptions concerning the etiologic role of an observed property that must be introduced into the analysis if the observations are to be taken as support for any specific claim. The complete set of background assumptions (premises) required to infer a given claim from a given observation is referred to as the relevance of the observation statement to the task at hand. Relevance of an observation then increases as support for the necessary set of background assumptions grows stronger. A set of observations and the associated judgment of its relevance, taken as a whole, is referred to as the warrant (Toulmin, 1958) for a claim to carcinogenicity. The strength of the warrant for a given claim increases with (1) the support of the observational base, (2) the strength of the background assumptions needed for relevance increases and (3) the degree to which the analyst explicitly recognizes the necessary background assumptions and incorporates them into discourse and reflection. The last requirement arises from the idea that premises must not only be true, but must be recognized as such if an analyst is to make a claim to rationality (Alston, 1985).

Since the risk from carcinogens arises from an interaction between the carcinogen and a biological system, the end of a complete hazard identification is to depict components of both the carcinogen (such as structure, physical state or concentration) and the biological system (such as physiological properties and mitotic rates) contributing to the existence of a hazard.

This end is formalized in the potential claims to carcinogenicity noted previously. By subdividing the claims into categories related to mode of action and context within which the action takes place, the analyst aids the end of risk mitigation discussed earlier.

In summary, the current report assumes that a rational framework for hazard identification (and, hence, an "informed judgment" of carcinogenicity) requires the following components:

(1) Consideration of the total body of evidence to be used in warranting depictions of risk. For example, the evidence might be data from in-vitro assays, epidemiologic studies, measurements of mitotic rates, etc.

(2) A judgment that the evidence invoked by the analyst is an adequate representation of the full body of available evidence (i.e. constitutes the "hard look" as discussed in the second interim report).

(3) Assessment of the foundational quality (see the appendix) of each separate piece of evidence used in the analysis. In other words, this is a determination of whether a given piece of evidence is considered reliable as an observational claim (an issue separate from that of the relevance of the evidence).

(4) Reflection on the relevance of each piece of evidence to each depiction of risk. This includes consideration of all relevance strategies (such as scientific theories constituted by etiologic premises) hypothesizing a role played by the measured factor in the production of risk (here, the risk of cancer).

(5) Judgment of the epistemic status of a given claim of carcinogenicity. This status reflects the foundational quality of the evidence; the relevance of that evidence to the stated task as specified by each relevance strategy; the degree of coherence between evidence that strengthens and weakens claims made in the analysis; the evidential support for any relevance strategy and its body of background premises used in the analysis; and philosophical reflection on the nature of evidence and the relationship between this nature and assignment of epistemic status.

1.2. The Need for Rational Discourse

Formal (mathematical) decision rules (Haseman, 1990; Eddy, 1989; Eddy et al.,1990b) may be helpful at interim steps in assembling and assessing observational evidence, but such approaches do not eliminate the need for human judgment and discourse for rational decision

making. First, the concept of probability typically employed in mathematical approaches can be formalized best by consideration of a single piece of evidence (or homogenous body of evidence) and its relationship to a single claim. While such considerations are useful in judging the foundational epistemic status of each single piece of observational evidence (such as the probability that the true mean of an underlying sampled population is X), probabilistic approaches are not well developed for instances of judging the coherence between multiple, dissimilar and potentially incommensurable bodies of evidence. Both foundational and coherence issues underlie scientific discourse and justification (Laudan, 1977). The second concern with relying on formal mathematical tools is the contention of the authors that rational judgments neither require an algorithm of belief (Brown, 1988) nor are properly summarized through mathematical probabilities.

A key assumption is that rationality is linked most intimately to discourse concerning epistemic status, with this discourse being guided by (but not formalized by) principles of evidential reason. As stated by Bernstein (1983), "central to this new understanding is a dialogical model of rationality that stresses the practical, communal, character of this rationality in which there is choice, deliberation, interpretation, judicious weighing and application of universal criteria, and even rational disagreement about which criteria are relevant and most important." This discourse is typified by the deliberations of a Science Advisory Board Committee and the attending parties. Such discourse does not prevent mathematical formalization when it captures the full quality of the discourse, but the burden is on formal tools to display their utility rather than on the discourse to fit into the axiomatic base of the formal tools. The intent of this report is to present an integrative framework for discourse on rational justification of hazard identification against which all formal tools may (if so desired) be compared.

To summarize, this research is concerned with the epistemic link between evidence, relevance strategies and claims about carcinogenicity, as well as the rational basis for discourse concerning this link. The term "claim" is chosen deliberately to distinguish it from "belief". Having provided the claim and its associated epistemic status, there remains an important issue as to when a claim, characterized by a given warrant, is to be elevated to the status of a belief by some person or group. This latter issue, while important in completely justifying the assertion that beliefs of an agency are rooted in rationality, is not addressed here since it involves components of psychology, sociology and political responsibility falling outside the domain of the present research. The present document focuses, therefore, on the evidential support for a claim (such as the claim that a substance is or is not a carcinogen), but not on the process by which claims are translated into beliefs. This presumes that it is possible to speak of epistemic status in a non-normative manner (i.e. without reference to the manner in which epistemic status should be related to the adoption of a belief). Complete rationality in policy matters, of course, requires that this latter issue of translation be discussed by any "users" of the claims generated in the risk analysis.

1.3. Elements of the Framework

A claim regarding carcinogenicity of an agent is the product of human judgment applied to evaluation of the informational base. There are, however, multiple sources of observational evidence, claims to carcinogenicity, and strategies of relevance for linking the two. The procedure to be described in this report was developed around certain principles and premises, such as completeness of evidence relevant to hazard identification, incorporation of qualitative as well as quantitative characteristics of evidence, a rational basis for warranting claims of

carcinogenicity from evidence, assessment of coherence of claims, and the centrality of informed human judgment to decision-making.

These characteristics are derived from consideration of the epistemic status of a depiction of risk (here, a claim of carcinogenicity), which is related to both the observational data available to the analyst and the relevance of these data to each of the specific taxonomic claims of carcinogenicity. This suggests that the epistemic status of a claim of carcinogenicity is determined by three judgments related to each piece of observational information:

(1) What is it that has been observed and how well has this observation been established? This judgment encompasses the first three of the five framework components listed in Section 1.1.

(2) What is the relevance of the observation to each specific claim? This judgment is given by the fourth component listed in Section 1.1.

(3) In what way does this observation support and/or detract from the claim that the substance is a carcinogen in any sense? This judgment is given by the fifth component listed in Section 1.1.

The first issue is related to the attempt to define the observational base on which any claims made in the analysis will (in some manner) be warranted. The analyst must establish the degree to which each claim to observation is considered well founded, regardless of the use to which the observation subsequently is put in the analysis. In the terms of logical positivism, the intent here is to produce "observation sentences" (Newton-Smith, 1981) from which the claims of the final analysis may be constructed when employed in conjunction with background premises required for relevance. These observation sentences should encompass all observations available on a given substance (more on this below). For simplicity in this report, the term "observation sentence" is replaced by the term "observation", although the two terms have distinctly different meanings in epistemology.

The second issue is related to the definition of relevance of the observations to the task of warranting a claim of carcinogenicity (i.e. that the substance is a carcinogen in some sense). As described in the appendix, observations gain relevance through the use of background premises relating the observation to a claim of carcinogenicity. For example, observation of DNA adducts gains relevance through the premise that adducts are a first step in the process of neoplastic conversion (Perera, 1987; Craig et al., 1981; Belinsky et al., 1987c), or are correlated with conversion (examples of relevance strategies, formalized in Section 5.1.). While multiple observations may be used to warrant a given claim, it also is the case that multiple sets of background assumptions may be used to assign relevance to a given observation. The analyst must, therefore, establish the base of strategies for relevance by which observations are folded into the analysis.

The framework to be described in the following chapters can be implemented on a substance of interest by following the seven steps for its application given in Chapter 6. To develop the framework, however, we need to consider: (1) determination of a taxonomy of claims of carcinogenicity, (2) the collection and assessment (both qualitative and quantitative) of the strength of observational evidence, (3) the assessment of the relevance and coherence of evidence for warranting claims of carcinogenicity within the observational context of that evidence, i.e. under the actual conditions of observation, particularly with respect to species and exposure level for data on humans or animals ("context" is formerly defined in Section 4.1.), and (4) the same as (3) except with extrapolation from one context to another, such as exposure to humans instead of animals or exposure at environmental levels instead of at higher concentrations. Stated briefly, we need to address four basic questions: What is meant by "carcinogenicity"? What observational data should be initially collected and assessed for their utility in judging carcinogenicity? How should the evidence be interpreted for hazard

identification within the observed context (i.e., without extrapolation across doses or across species, assuming the availability of data from at least one epidemiologic study or long-term animal study)? What interpretation can be made with extrapolation (from high exposure to low exposure or from exposure of animals to exposure of humans)?

These four questions are addressed in Chapters 3-5, following discussion of some fundamentals regarding cancer mechanisms and a "biologically significant dose rate" in the next chapter. The broad theoretical discussion of Chapter 2 constitutes a "minimal" set of etiologic assumptions to be imposed on the analysis. This minimal set provides structure without forcing judgments into the boundaries of a particular theory of carcinogenesis. In other words, existing theories tend to share this common axiomatic base, allowing the resulting discourse to remain fairly "theory-neutral" until explicit relevance strategies are invoked. It should be recognized, however, that the existence of any "theory-neutral" framework is highly controversial in the literature on epistemology and philosophy of science (the classic text here is Kuhn, 1962).

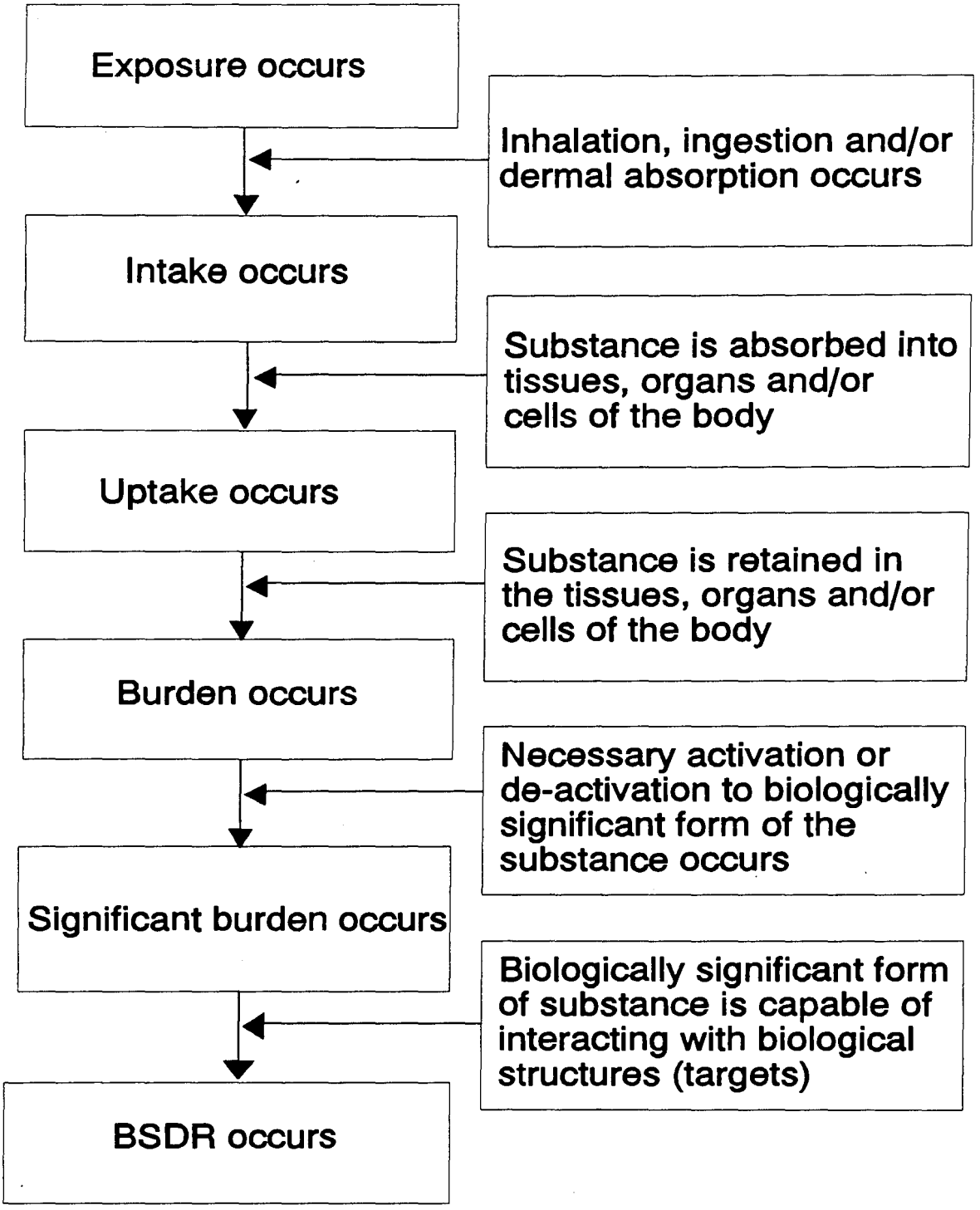
2. CANCER MECHANISMS AND BIOLOGICALLY SIGNIFICANT DOSE

Two central concepts employed in causal theories of carcinogenesis are those of (1) a biologically significant dose-rate (BSDR) within the body (Andersen, 1989; Gerlowski and Jain, 1983; Lutz and Dedrick, 1987) and (2) transitions between states of cancer brought about by the BSDR (Moolgavkar, 1991). The BSDR is produced through a series of physical steps related to the following conceptual categories (see Figure 1):

(1) Exposure to the substance in the environment produces an intake into the body by various routes (via the lungs, G.I. tract, or skin).

(2) The intake results in an uptake into the body, if the substance is deposited in the body following intake. This distinction between intake and uptake is required by the fact that substances can, for example, be inhaled without depositing in the lung (such as in the case of inert gases).

Figure 1. Flow Diagram of Judgments for BSDR



(3) The uptake, in conjunction with retention of the substance in the body, yields a burden (i.e. concentration of the substance in the body). Increases in either uptake or retention generally yield increases in burden. The distinction between uptake and burden is important because a substance may be taken into the body but immediately removed, not allowing time for interaction with whatever tissue, cell, organ, etc., constitutes the target for the effect. In addition, etiologic theories of carcinogenesis may posit a threshold burden below which cancers do not appear even if an uptake occurs.

(4) Burden refers to the amount of the original substance present in the body. For many carcinogens, biotransformation (metabolism) changes the chemical form of the original substance (Whitey, 1982; Vainio and Hietanen, 1980; Clayson, 1985; Gehring et al., 1978; Farber, 1987; Miller and Miller, 1976). This new chemical form may be either (1) the form responsible for cancer, in which case the biotransformation results in activation or (2) a form incapable of inducing cancer, in which case the biotransformation results in deactivation. In any event, theories of carcinogenesis presuppose that there are some chemical or physical forms of a substance capable of inducing cancer and other forms incapable of inducing cancer. The former is termed here the biologically significant form. The burden of this biologically significant form of the original substance in the body is termed here the biologically significant burden. The distinction between this B-S-B and the burden is important because a substance may be present in the body without necessarily being present in the form capable of inducing cancer. This can occur if enzyme systems required for activation are not present or if systems required for deactivation transform all of the biologically significant substance into an inactive form.

(5) Finally, the biologically significant burden (BSB) produces a biologically significant dose rate (BSDR) if the biologically significant form of the substance is capable of interacting with biological structures in the body (DNA, membranes, etc.) believed to be the site of action for carcinogenesis (Andersen, 1989; Barrett and Wiseman, 1987; Farber, 1987). The distinction between BSB and BSDR is important because a substance may be present in the body but unable to interact due to the presence of barriers to interaction or a lack of important sites of interaction.

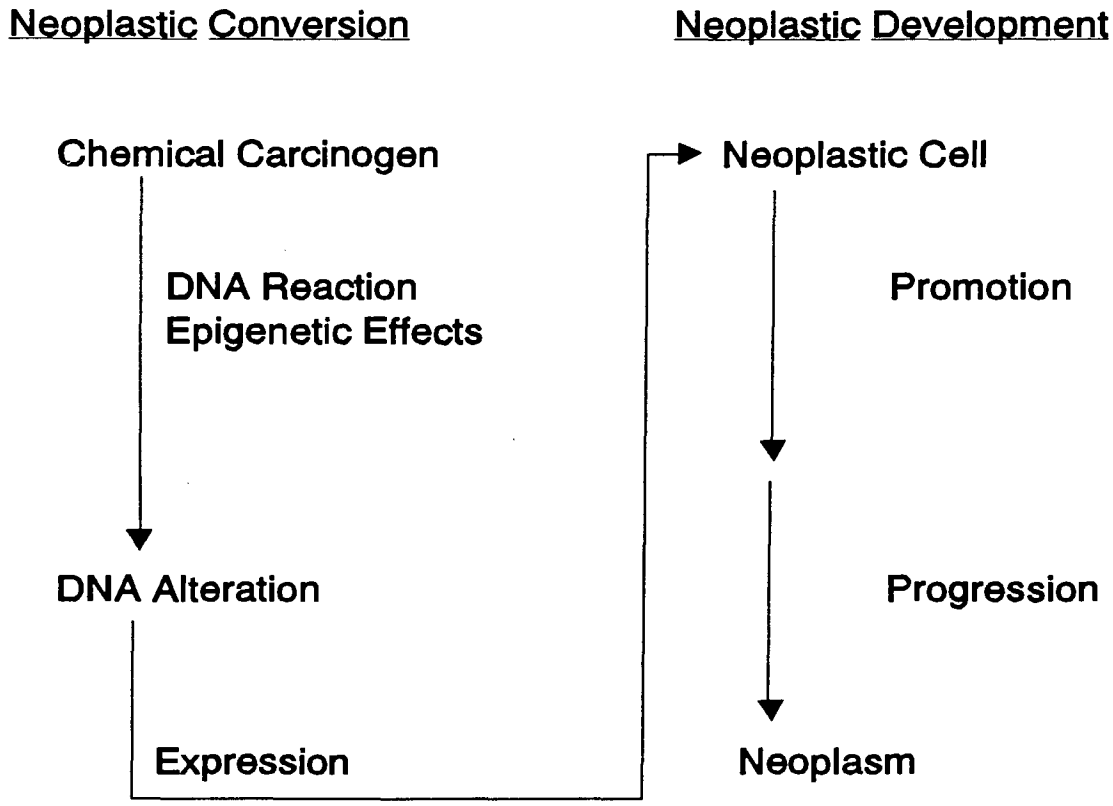
The above five considerations of the process leading to a biologically significant dose-rate introduce premises (to be discussed in more detail in Chapter 5) necessary in ensuring that exposures in the context of interest to the analyst result in at least some BSDR. We turn now to the issue of changes in states of cancer as induced by this BSDR. In the broadest and least theoretical sense, it might simply be premised that exposure to a substance (and, hence, production of a BSDR) induces a change from "normal" health to cancer. The probability of

this change presumably depends upon a background rate of change, the BSDR, and the length of time over which the BSDR is maintained in the body. A more general way of stating this is that the change depends upon the temporal pattern of the BSDR over the lifetime of an exposed organism.

Current theories of carcinogenesis differentiate between a number of stages or states (by "state" here, we do not necessarily mean a morphologically, physiologically, etc., identifiable condition; different theories of carcinogenesis identify states differently and may even drop the concept of a state) intermediate between "normal" and cancerous. It generally is accepted that these states consist of at least neoplastic conversion and neoplastic development as depicted in Figure 2 (Williams and Weisburger, 1991). Cells are presumed to pass from normal to converted and finally from converted cells to developed tumors. Only cellular clusters developing to a frank tumor possess the capability of inducing fatality as a result of the cancer. Other cells or cellular clusters are considered to be precancerous. Changes from one state to another typically are referred to loosely as transitions. (By "transition" here, we do not necessarily mean a distinct stochastic change of state, but rather a general historical movement towards one of the states of cancer; different theories of carcinogenesis treat transitions differently, with individual theories focusing on transitions as stochastic events, alterations in kinetics, etc.). Identification of the transitions as initiation, promotion and progression are common in the literature (Pitot and Campbell, 1987; Barrett and Wiseman, 1987) but these terms are more operational than mechanistically descriptive and are not included here as part of the "minimal" theoretical base.

The above discussion of theories of carcinogenesis focuses on changes or transitions without giving explicit mechanisms for those changes. While there is general agreement that cancer is a multi-stage phenomenon, there is less (although still substantial) agreement that the

Figure 2. Main Steps in the Carcinogenic Process



Adapted from Figure 5-1 of Williams and Weisburger (1991).

stages should be identified as neoplastic conversion and neoplastic development. There is even less agreement as to the mechanisms by which these stages arise and whether there are mechanisms common to carcinogens.

In the case of neoplastic conversion, the dominant etiologic theory relates conversion either to changes in DNA (Lawley, 1987; Farber, 1987; Williams and Weisburger, 1991), referred to as genotoxicity, or to epigenetic changes (Bartsch and Malaveille, 1990; Barrett, 1987; Butterworth, 1990; Perera, 1984), referred to (more generally) as non-genotoxicity. The former changes are unclear at present, but candidates for important change are: (1) formation of DNA adducts (Belinsky et al., 1987c; Farber, 1987), (2) changes in base-pair sequence, (3) single or double-stranded DNA breaks (Slaga, 1988; Tennant et al., 1987a), (4) chromosomal aberrations (Lewis and Adams, 1987; Hall and Freyer, 1991), and (v) activation of oncogenes (including deactivation of repressor genes) (Stowers et al., 1987; Aaronson and Tronick, 1986; Bos, 1988). Associated with the above four mechanisms are concepts of misrepair (through improper insertion of bases) of adducts and DNA breaks (Curtis, 1991) and translocation of broken DNA to allow expression of oncogenes (Della-Favera et al., 1988; IARC, 1986). Epigenetic (non-genotoxic) changes refer to alterations in cellular structures other than DNA, such as membranes and antigens, which might, in turn, yield changes to DNA (although DNA changes are not necessarily part of neoplastic conversion).

In the case of neoplastic development, mechanistic premises have tended to focus on intercellular communication (Trosko and Chang, 1988; Harper and Legator, 1987; Langenbach et al., 1988; Slaga, 1984), hormonal control (Moolgavkar, 1986; Williams, 1990) and the kinetics of growth and removal within organized cellular communities (Homburger and Treiger, 1969; Loury et al., 1987; Swenberg and Short, 1987; Cohen and Ellwein, 1990; Short and Swenberg, 1990). Changes in these properties might arise from either genotoxic or epigenetic action,

although the latter usually is cited as the route of action. There is little agreement as to whether these changes are potentially reversible although it has been suggested that promotion has both a reversible (Phase I) and irreversible (Phase II) component (Barrett and Wiseman, 1987; Hermo and Brandt-Rauf, 1987; Langenbach et al., 1988). As in neoplastic conversion, the largest source of disagreement is over premises concerning mechanisms of action. Candidates are (1) changes in DNA (Bailey et al., 1991), making neoplastic development similar to conversion, although this aspect of promotion probably is limited to Phase II; (2) interference with intercellular or intracellular communication through changes in messenger molecules, gap junctions, microtubule structure, etc. (Trosko and Chang, 1988); (3) disruption of the histological architecture of cellular communities as in the theory by (Tamplin and Cochran, 1974); (4) induction of hyperplasia in the cellular community through stimulated mitosis or other (currently unspecified) means (Ames and Gold, 1991a; Schulte-Hermann et al., 1991); and (v) changes in the action of hormones on cells through changes in the amount of hormone present at the target organ, the form of the hormone, the density of receptors for hormones and/or the specificity of hormone receptors (Moolgavkar, 1986).

3. MEANING OF "CARCINOGENICITY"

3.1. Current Classifications of Carcinogenicity

The criteria for chemical group classifications currently are ranked according to the supporting evidence on changes in tumor prevalence in humans and animals, with human evidence considered more relevant. For example, in the classification system used by the EPA, the taxonomy is: Group A: Human carcinogen, with sufficient evidence from epidemiologic studies; Group B1: Probable human carcinogen, with limited evidence from epidemiologic studies; Group B2: Probable human carcinogen, with sufficient evidence from animal studies and

inadequate evidence or no data from epidemiologic studies; Group C: Possible human carcinogen, with limited evidence from animal studies in the absence of human data; etc. (U.S. EPA, 1986). Classification schemes of the International Agency for Research on Cancer (IARC), the National Toxicology Program (NTP), and the American Conference of Governmental Industrial Hygienists (ACGIH) also have this characteristic; i.e., at the upper end, the strength of the claims is dependent on the extent of evidence from human epidemiologic studies, proceeding to weaker claims associated with evidence of changes in tumor prevalence from animal studies. The single exception is an "indirect" classification of ACGIH which refers to carcinogenicity activity that occurs primarily as secondary effects of some other toxic or physiological action by the substance or its metabolites (Cohrssen and Covello, 1989), although even here the evidence for such activity is change in tumor prevalence.

Several properties in common to these agency's classification criteria may be noted: (1) The strength of the claim to human carcinogenicity is limited by the kinds of data available rather than reflection upon the complete body of data on a suspect substance. e.g., an agent could not be classified strongly as a "human carcinogen" (Group A or B) by EPA if there were no epidemiologic data available, regardless of how convincing a case might be made on the basis of other evidence and scientifically warranted lines of reasoning. This suggests the existence, at least implicitly, of some conception of "minimal epistemic status" (see the discussion in the appendix) associated with specific warrants. (2) The criteria do not state, but give the impression, that if available in adequate quantity and quality, then animal data alone are sufficient for classification of a substance as a "probable human carcinogen". The manner in which animal data play such a role, however, is not specified, nor are the criteria specified by which premises necessary for inter-species extrapolation are to be warranted. (3) "Human carcinogen" suggests that if humans differ in susceptibility to a carcinogen, it is only by degree,

i.e., that it is not possible for an agent to be a human carcinogen only to a subpopulation that may be predisposed in some way. (4) There is no reference to the context within which a substance acts as a carcinogen, requiring the (perhaps unwarranted) premise that carcinogenicity is a property of the exposure factor alone rather than an interaction between exposure to the factor of interest, biological properties of the exposed organism, and concurrent exposures. These four features of existing classification schemes presume that a substance may be categorized only as a "carcinogen" with epistemically-based modifiers (possible, probable, etc.) depending upon the category of evidence available.

With its focus on carcinogenicity as a single claim and on epidemiologic and whole-animal data, the existing classification scheme does not make explicit the etiologic differences which underlay various carcinogens and the manner in which data other than whole-animal carcinogenicity assays or epidemiologic studies may (1) elucidate these differences and (2) strengthen or weaken claims to carcinogenicity within a given context. The framework of analysis developed here extends the existing procedures for hazard identification to include the above considerations.

3.2. A Taxonomy of Claims of Carcinogenicity

In the present report, three assumptions are made for a claim of carcinogenicity within a framework for rational discourse on hazard identification: (1) the claim is manifold (i.e. has different taxonomic forms useful for meeting different ends of risk mitigation); (2) the claim is dependent on the nature and strength of the full body of supporting evidence on the agent considered; (3) the claim should be formulated uniquely in each specific case. This suggests that hazard identification for carcinogens might best be served by a taxonomic scheme giving explicit recognition to the role of a substance in carcinogenesis and, hence, the context under which the

substance increases the prevalence of cancer. Implicit to any claim of carcinogenicity is a context, e.g., long-term animal exposures, humans exposed in an occupational setting, etc. Consequently, the taxonomic scheme is applied to an agent/context combination, even when explicit reference to the specific context is omitted. The uncertainty in claims of carcinogenicity then will be a function of the context towards which those claims are directed.

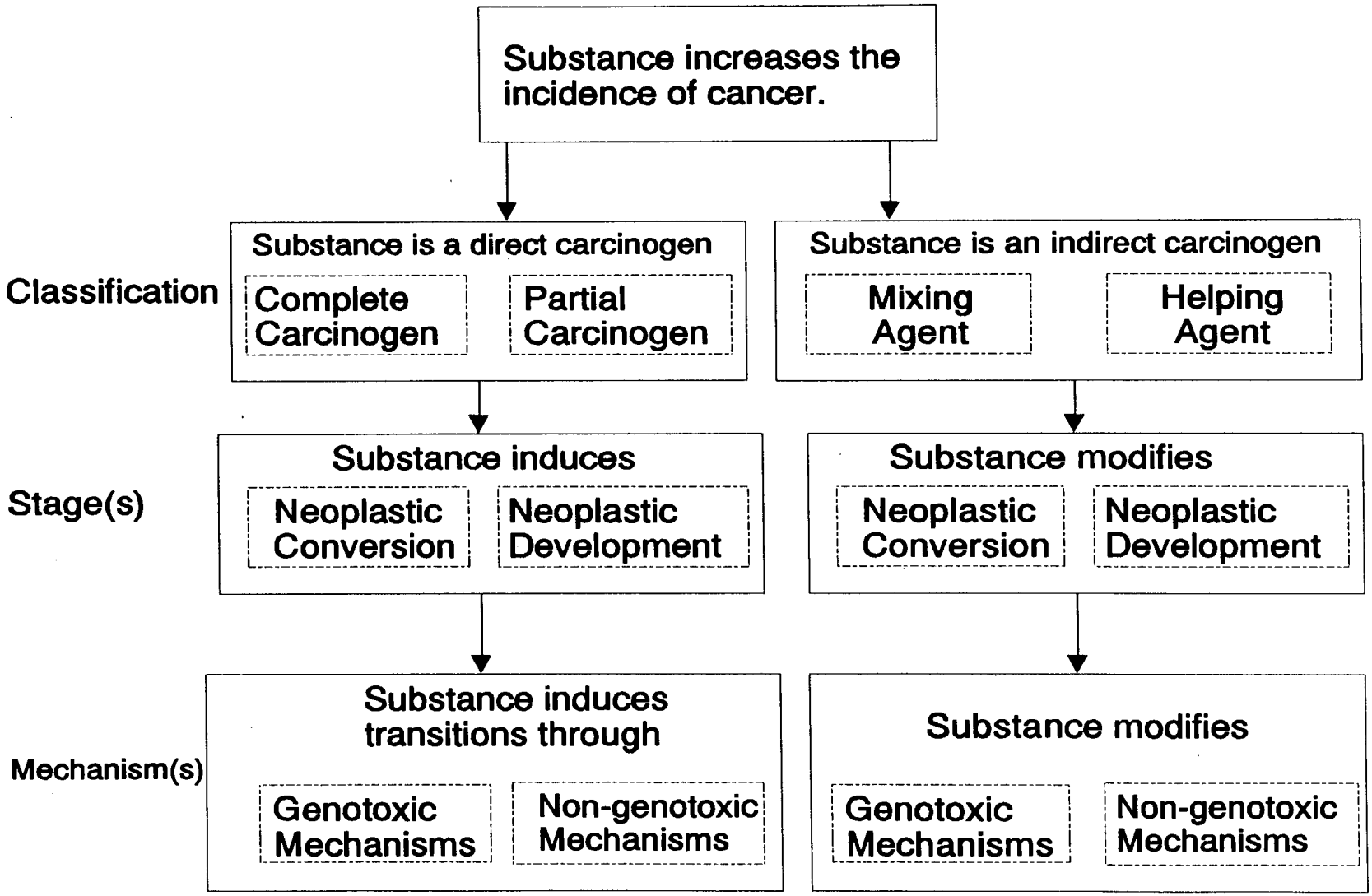
The taxonomy employed in the present report is described in Table 1 and is depicted in Figure 3. At the first and least informative level, the risk analyst might make the claim simply that a substance increases the incidence of cancer (within a specified context). At the second level, the analyst might make the claim that the substance is either a *direct carcinogen* (*partial* or *complete*) or *indirect carcinogen* (*mixing agent* or *helping agent*). These terms are defined in Table 1, which also describes conditions under which the claim would be preserved in another context. The distinctions between "direct carcinogen", "mixing agent (mixer)", and "helping agent (helper)" are important. A direct carcinogen (the parent agent or a biologically active product) induces one or more transitions between states or "stages" in the cancer process by itself, without requiring the presence of another chemical agent. If it induces all the necessary transitions leading to a neoplasm, then it is called complete; otherwise it is called partial. For example, formaldehyde, radiation, BaP, and most other substances that are commonly called carcinogenic are complete carcinogens. Some examples of incomplete carcinogens are urethane, which is an effective initiator but not a promotor (Berenblum and Haran-Ghera, 1957; Barrett and Wiseman, 1987); phorbol esters, which are active promoters without initiating activity (Boutwell, 1974; Barrett and Wiseman, 1987); and probably arsenic, which appears to act primarily to increase progression (Barrett, 1984). A substance is a mixing agent if it is not a direct carcinogen, but a direct carcinogen is produced when it is mixed with one or more suitably chosen substances that also is not a direct carcinogen, i.e., if a mixture of substances,

none of which is a direct carcinogen itself, interact biologically or chemically to effect transitions, then the substances are called mixing agents. Note that by definition a mixing agent must be able to produce a direct carcinogen from interaction with some other mixing agent, not with a substance that is already a direct carcinogen (either complete or partial). If a substance is not a direct carcinogen or a mixing agent, but it enhances one or more transitions induced by a suitably chosen direct carcinogen, then it is called a helping agent. A helping agent is sometimes referred to as a "risk modifier" in the literature. For examples of

TABLE 1. CLASSIFICATION SCHEME UTILIZED IN THE PRESENT REPORT

Classification	Definition	Antecedent Conditions Under Which Substance Elevates Cancer Incidence
Complete Carcinogen	The substance, acting directly on components of the organism, induces all changes required to elevate the incidence of cancer in a population.	Any conditions in which the mechanism by which the substance exerts its effect is present and operational.
Partial Carcinogen	The substance, acting directly on components of the organism, induces part of, but not all of, the changes required to elevate the incidence of cancer in a population.	Any conditions in which (i) the mechanism by which the substance exerts its (partial) effect is present and operational and (ii) the remaining changes required for cancer are produced in the population by other means.
Mixer	The substance, only when acting in conjunction with a second (mixer) substance, acts directly on components of the organism to induce part of (partial mixer) or all of (complete mixer) the changes required to elevate the incidence of cancer in a population.	Any conditions in which (i) the necessary second substance is present, (ii) the mechanism by which the two substances act in conjunction to exert their effect is present and operational and (iii), for partial mixers only, the remaining changes required for cancer are produced in the population by other means.
Helper	The substance, while not producing any of the changes required to induce cancer, produces changes in the antecedent conditions under which a second (carcinogen) substance exerts its effect, thereby magnifying the elevation in cancer incidence caused by the second substance.	Any conditions in which (i) the second necessary substance is present, (ii) the mechanism by which the original substance exerts its magnifying effect is present and operational, and (iii) the mechanism by which the second substance exerts its (carcinogenic) effect is present and operational.

Figure 3. Levels of Taxonomic Claims of Carcinogenicity



helping agents, with a limited discussion of possible mixing actions, the reader should consult Williams (1984), Berenblum (1985), Weisburger (1988), Hecker (1984) and Homburger and Treiger (1969).

Returning to the description of the taxonomy, at the third level in the taxonomy the analyst might differentiate between effects on neoplastic conversion, neoplastic development, or both steps to cancer. At the fourth level, the analyst might make the claim that the substance induces alterations in the organism through genotoxic and/or non-genotoxic routes of action.

4. INFORMATIONAL BASE

4.1. Assembling the Observational Base

Having established the taxonomy of claims that constitute the end of an analysis, the initial step in analysis is establishment of the base of observations on which warrants of a claim to carcinogenicity might be developed. This step establishes, to the extent possible, the existential content on which all parties engaged in an analysis (or debate concerning an analysis) might agree regardless of differences in interpretation of the observations. The existence of such a "theory-neutral" base is controversial (Longino, 1990), but the step is included here since there is substantial agreement within the field of carcinogenesis concerning the importance of a number of specific observations to claims of carcinogenicity (even if the interpretation of the observations varies between analysts). The present section describes the sources of information potentially available to an analyst when faced with the task of hazard identification. The collection and preliminary evaluation of all observational data of potential value in assessing carcinogenicity of a specific agent is a major undertaking. To facilitate organization and application in the suggested framework developed here, observational information is classified by data category and data item, as illustrated in Table 2. A data category refers to a set of

observations characterized by a common use in lines of reasoning (such as supporting a common background premise). A data item refers to a specific observation within the data category.

The category "Related Substances Assessments" refers to hazard identification of other agents, including mixtures, that might have some predictive value for the agent of interest through similarities in the etiologic role in carcinogenesis. For example, ETS is largely the product of sidestream smoke produced by a smoldering cigarette between "puffs". Sidestream smoke has been shown to be qualitatively similar in chemical constituency to the mainstream smoke that smokers inhale, containing several known or suspected carcinogens (U.S. EPA, 1990). Additionally, observational evidence strongly supports the conclusion that smoking increases the risk of lung cancer. Thus, for hazard identification of ETS, prior assessments of mainstream smoke provide information (from a related substance) useful in the line of reasoning referred to in the literature on ETS as "cigarette equivalence" (Thorslund, 1990).

The broad theoretical implications of the data categories in Table 2 are described in Table 3. The categorization shown in Table 2 was constructed to group data items supporting similar lines of evidential reasoning for claims of carcinogenicity (to be discussed in Chapter 5). Assessing the evidential support for carcinogenicity must also take into account (1) the context in which data were observed (e.g., a long term animal study on mice or an epidemiologic study of workers exposed at high concentrations) and (2) the context in which hazard identification is of interest (e.g., persons exposed at typical environmental levels). This raises the issue of what is to be meant by a context. In general, a context refers to a set of characteristics of an exposure situation believed to affect the carcinogenicity of the substance (or exposure factor) of interest. More formally, a context is an exposure-response scenario with specification of factors present that may affect the assessment of hazard. Observations within a given context then are assumed to share a common etiologic link between exposure and response, so that each

TABLE 2. SUMMARY OF THE ROLE OF OBSERVATION STATEMENTS IN EVIDENTIAL REASONING FOR CARCINOGENICITY CLAIMS

Data Category[1]	Data Item[1]	References Citing Role of Data Item	Specific Judgment[3]
Tumor Response[2]	Incidence or Prevelance	15 120 221 247 248	C
	Dose-Response	2 3 15 66 120 166 215 233 252	C
	Time-to-Appearance	24 120	C
	Multiplicity	120	C
	Age-at-Appearance	120	C
	Initiation	2 17 31 100 101 140 211 293	S
	Promotion	2 17 30 31 53 67 100 211	S
Biophysical Effects	Hyperplasia	6 8 66 173 175 247 248 252	PT
	DNA Adducts	18 22 34 42 63 64 68 69 241 247	PT
	Oncogene Activation	2 22 31 52 124	PT/T
	Interference with Intercellular Communication	32 125 126 169 231 261	PT/T
	Degree of Metastasis	81	T
	Concentration of Tumor Growth Factor (TGF)	50	PT
	DNA Breakage	54 88 100 101 124 173 250 289	PT
	Chromosomal Aberrations	54 69 88 100 101 124 250	PT/T
	Site-Specific Mutation	9 69 88 100 101 124 250	PT/T
	Mutagenicity	9 69 88 101 107 124 202 214 226	PT/T
	Cellular Transformation	33 88 100 101 124	PT/T/S
	Alterations in Membrane Permeability	230	PT
	RNA Alterations	171	PT
	Appearance of Cancer Marker Proteins	81 176 215 284	T/S/C
	Alterations in Cellular Antigens	81	T/S
	Presence of Preneoplastic Lesions	29 34 151 231 284	T/S
	Alterations of Cellular Architecture	254	PT
	Alterations of Distribution in Differentiation and/or Histology	103	PT/T/S
	Cytotoxicity	8 67 85 124 224 232 247 252 253	PT
	Tumors Appearing in Hosts after Injection of Transformed Cells	81	S/C
	Hormonal Alterations:	67	
	Hormone Production Rate	67	PT
	Hormone Structural Form	67	PT
	Hormone Binding Sites		PT

(continued on following page)

Table 2. (continued)

Data Category[1]	Data Item[1]	References Citing Role of Data Item	Specific Judgment[3]
Host Characteristics	DNA Repair Rates	85 88 96 223 241	PT
	DNA Repair Specificity	223	PT
	Density of DNA Repair Enzymes	223	PT
	Repair Kinetics	85 88 146 223	PT
	Activation/Inactivation of Repair Process	223	PT
	Background Transition Rates	35 67 85 134	PT
	Initial State Vector	67 85 134	PT
	Presence of Target Organ	35 96 120	PT
	Presence of Carcinogenicity Mechanism	35 96	PT
Pharmaco-dynamics	Rate of Inhalation:		
	Tidal Volume	3 11 105 106 122 182 208 222 247	EI
	Minute Volume	3 11 105 106 122 182 208 222 247	EI
	Rate of Ingestion:		
	Food	82	EI
	Water	15 197	EI
	Airway Diameters and Lengths	105 106 115 200	IU
	Airway Branching Scheme	105 115 200	IU
	Deposition Fractions	115 200	IU
	Epithelial Integrity	7 149 183 245	IU
	Air:Blood Partition	13 15 16 158 183 229	IU
	Integrity of Mucus Layer	247	UB
	Mucous Flow Rate	89 191 247 248	UB
	Peristaltic Velocity	182 197	UB
	Facilitated Transport:		
	Carrier Density	137 197	IU
	Binding Coefficient	137 197	IU
	Dumping Coefficient	137 197	IU
	Specificity of Carrier	137 197	IU
	Saturability of Carrier	137 197	IU
	First Pass Excretion	18 281	UB/BSB
	First Pass Metabolism	281	BSB
	Water:Oil Partition	197 208	IU
	Water:Lipid Partition	197 208	IU
	Cardiac Output, Q	21 138 183 208 222	IU
	Organ Perfusion	21 144 197	IU
	Pore Size (membranes)	149	IU

(continued on following page)

Table 2. (continued)

Data Category[1]	Data Item[1]	References Citing Role of Data Item	Specific Judgment[3]
Pharmaco-dynamics (continued)	Facilitated Transport Energy Source	137 149	IU
	K_M for Metabolic Reaction	58 93 108 138 154 208 220 221 239 274	BSB
	V_{max} for Metabolic Reaction	38 58 59 93 137 138 208 220 221	BSB
	Substrate Density	37 38 39 78 121	BSB
	Presence of Competing Metabolic Pathways	58 142 172 155 287	BSB
	Renal Flow Rate	242 273	IU
	Permeability of Renal Tubes	242	IU
	Excretion Rates	18 161 221 242 273	UB
	Identification of Active Metabolite	20 37 49 93 95 120 206 220	BSB/BSDR
	Identification of Target Cells	120 252	IU/BSB/BSDR
	Identification of Site of Metabolism	12 37 112 162 287	BSB
	Substance Diffusion Coefficient	149	IU/UB
	Neutrophilicity	187	BSDR
	Adduct Binding Coefficient	108 159	BSDR
	Enzyme Concentration	112	BSB
	Intake	15 65	EI
	Uptake Fraction	15 65 245	IU/EI
	Burden	15 65 225 245 246	UB/IU/EI
	Dose Rate or Dose	15 246	UB/IU/EI
	Biologically Significant Burden	15 42	BSB/UB/IU/EI
	BSD or BSDR	15 42 159 160	BSDR/BSB/UB/IU/EI/UB
	Retention Function	15 222	UB
Concurrent Environmental Conditions	State of Attachment to Other Substances	200 288	EI/IU/EB
	Presence of Oils in Administered Dose	288	IU
	Presence of Other Substances in the Environment	79 114 130 156 194 215 228	PT
Structure Activity Relationships	Bay Region Site	25 26 27 83 120 192 256	BSDR

(continued on following page)

Table 2. (continued)

Data Category[1]	Data Item[1]	References Citing Role of Data Item	Specific Judgment[3]
Related Substances Assessments		see, e.g. 259	ALL
Environmental /Substance Characteristics	Particle Size	200 228	IU
	Particle/Fiber Shape	200 228	IU
	Particle Hygroscopicity	200	IU
	Concentration Variability		
	Spatial	84	EI
	Temporal	84	EI
	Chemical Form		ALL

1 See Working Table 2.

2 Quantitative measures are shown under "data items" for "tumor response" only. Quantitative measures commonly used are not given in other data categories, only the endpoints of interest.

3 This refers to one or more of the specific judgments that must be made in using Figures 1 and/or 2 for supporting claims of carcinogenicity through theory-based inference. The key is as follows:

EI:	conversion from exposure to intake
IU:	conversion from intake to uptake
UB:	conversion from uptake to burden
BSB:	conversion from burden to biologically significant burden
BSDR:	conversion from biologically significant burden to biologically significant dose rate
S:	indicates a state or stage of carcinogenesis
T:	indicates a transition process between states or stages of carcinogenesis
C:	indicates carcinogenesis directly
PT:	partial transition; indicates an effect leading to (but not sufficient for) transitions
ALL:	affects all of the conversions.

TABLE 3. BROAD THEORETICAL IMPLICATIONS OF DATA CATEGROIES

Data Category[1]	Broad Theoretical Implications of Data Categories[2]
Tumor Response[2]	Direct observation of the effect of interest (cancer). Intra-context claim for carcinogenicity requires only the premise that any noted associations are causal.
Biophysical Effects	Provides evidence that a causal factor (such as exposure to the substance of interest) yields effects deemed important in the etiology of cancer. Since the observations are not of cancer directly, a theory of the etiology of cancer is required. Such theories would support the contention either that a given data item is (i) an indicator of cancer being present (e.g. tumor growth factor, appearance of cancer marker proteins, alterations in cellular antigens, degree of polyploidy), (ii) an indicator of transitions between states of cancer (e.g. oncogene activation, interference with cellular communication, mutagenicity, cellular transformation) or (iii) an indicator of conditions (mechanisms) necessary for transitions (remaining data items).
Host Characteristics	These data provide evidence that there is not a unique pre-existing characteristic of the organism that would preclude conclusions being drawn with respect to the carcinogenicity context in which the organism appears in the analysis. The intent is to ensure that an observed effect being associated with the BSDR cannot be attributed to some unique characteristic of the particular organism being observed. This category does not refer to the characteristics affecting exposure and BSDR.
Pharmaco-dynamics[3]	Provides evidence that the factor (such as exposure to the substance of interest) results in a biologically significant dose-rate to the organism. Since pharmacodynamic data do not constitute observations of effect, they do not provide either direct empirical or semi-empirical warrants for carcinogenicity claims. They can, however, provide conceptual support for the claim that any effects observed in the "Tumor Response" and "Biophysical Effects" data category are causally connected with exposure.
Concurrent Environmental Conditions	These data support the contention that there is nothing unique about the exposure conditions that would call into question inclusion of a given study into a given context, or extrapolation of an observed effect from one context to another. These conditions occur prior to intake of the substance and may affect the intake magnitude, intake route, uptake, etc. of the substance. These data also support the contention that an observed effect was (or was not) due entirely or partially to confounding exposures.
Structure Activity Relationships	Certain chemical/physical structures have been determined to be associated most strongly with carcinogenicity, particularly with respect to initiation. These structural features presumably are an indication of the ability of a substance to act on the organism by mechanisms governed through general topological and/or chemical properties of molecules.

1 See Working Table 2.

(continued on following page)

Table 3. (continued)

[2] Quantitative measures are shown under "data items" for "tumor response" only. Quantitative measures commonly used are not given in other data categories, only the endpoints of interest.

[3] Quantitative data on pharmacodynamics are typically utilized in the dose-response step of risk assessment, rather than in hazard identification. Pharmacodynamic data are included in Table 3, however, because they contribute qualitative information useful to the hazard identification step. Specifically, it is useful to know that exposure to the substance of interest results in a biologically significant dose-rate to the organism of interest if a claim of carcinogenicity is to be supported.

observation in the context may be taken as one sample from a potential population of samples sharing a common relationship between exposure and response. This stage of specifying contexts involves judgments founded in etiologic theories concerning the factors that might potentially exclude a given observational setting from a given context. In the language of philosophy of science, a context is defined by specifying the antecedent conditions under which a response is presumed to follow from exposure (see the appendix). It may even be the case that in-vitro data are included within a context otherwise defined by exposure to an animal or human, so long as the exposure-response characteristics of the in-vitro system are deemed to share an etiologic link with characteristics of the organisms in the context. This imprecise definition is intended to be more operational than descriptive, as the most one can do is to include within the specification of a context all factors considered to be of potential significance to the judgment of carcinogenicity.

Contexts are defined around actual or hypothetical exposures of animals or humans. If a context describes a scenario within which study data are available, typically from either an animal study (preferably, long-term) or epidemiologic study, then it is referred to as an observational context. A target context describes an "endpoint" of hazard identification, including a non-observational context requiring extrapolation of judgments of carcinogenicity across species or dose levels constituting the observational contexts. For example, suppose there are two sets of study data on substance X, one a controlled animal study and the other an epidemiologic study of persons exposed at atypically high levels, perhaps due to their location or workplace environment (as in the case of exposures to airborne radon (Cross, 1987) or formaldehyde (Graham et al., 1988)). A separate observational context may be defined for each study (since the two studies are judged not to share a common etiologic link) and perhaps a target context defined that is descriptive of typical environmental exposures. Alternatively one might have one or more animal studies with no epidemiologic studies, requiring that target contexts might be defined for human exposure (as in the case of acetaldehyde (Woutersen et al., 1986)).

Working Table 1 is provided for a description and reference number of each observational context and each target context.

Observational contexts may suffice without definition of a separate target context. For example, epidemiologic studies of ETS tend to be conducted under typical living conditions, descriptive of the context of concern for assessing hazard of lung cancer to adults or other endpoints, such as specified respiratory effects in children. Judgment needs to be exercised in defining contexts. As a rule of thumb, two contexts should differ in at least one characteristic of potential consequence for hazard identification (i.e. by at least one difference in important antecedent conditions). This is an imprecise procedure, on which judgments may differ. The conflict is between a need to summarize data for manageability and decision-making without sacrificing useful information in the process. (Technical note: This notion is somewhat analogous to the statistical concept of a sufficient statistic.) For example, there are at least 26 case-control studies on ETS and lung cancer. Depending on one's judgment of the homogeneity of the conditions under which the studies were conducted, from one to 26 contexts could be defined.

If studies within each country are considered sufficiently homogeneous to pool the results by methods of meta-analysis (DerSimonian and Laird, 1986; Greenland, 1987; Eddy, 1990; Eddy et al., 1990a,b), but the same is not the case for studies in different countries, then a separate context might be defined for each country. (Note: There are currently no firm guidelines on when, or even "if" in the minds of some analysts, meta-analysis is appropriate (Mann, 1990)). When there is more than one set of study data on animals or humans within a context, then the separate sets are termed cases. For example, if ETS studies are grouped into contexts by country (arguing essentially that subpopulations within a country experience similar antecedent conditions, but that the same is not true across

WORKING TABLE 1. CONTEXT SPECIFICATION FOR HAZARD IDENTIFICATION

Context Number (1, 2, ...)	Context Type (O,T)[1]	Description

[1] O: Observational Context
T: Target Context
Observational contexts and target contexts are numbered separately.

countries), then the individual studies within a country are cases. If studies between countries are considered homogenous, a single context might be defined. Similarly, long-term animal studies conducted by NTP generally include two species and two sexes. One to four contexts could be defined. If it is reasonable to combine results for the two sexes within a species, however, then one would have two contexts with two cases each. A context needs to include as many cases as reasonable so as to facilitate comparisons and judgments of coherence across studies, while still retaining as much specificity of detail as may be relevant to hazard identification. Data categories in Table 2 that are exclusively context-specific include: (1) tumor response, (2) biophysical effects, (3) concurrent environmental conditions, and (4) environmental/substance characteristics. The remaining data categories are largely species-specific or chemical-specific without reference to the context within which the given species is exposed to the chemical (although there will be cases in which even pharmacodynamic data will be context-dependent, as when absorption fractions, intake characteristics and/or retention characteristics change with magnitude of intake (for examples of formaldehyde see Graham et al., 1988; Burkhart et al., 1990)). Data items from these latter categories are included in Working Table 2 (to be completed for each observational context in Working Table 1), as applicable to the context. In other words, a data item (such as presence of the necessary bioactivation enzymes) is included in the working table for a given context even if the item has not been observed under the conditions of exposure defining the context, so long as it is judged that the item is species-specific (the species also defining the context). Any use of the data item to draw intra-context inferences of carcinogenicity will, of course, require introduction of the premise that exposure (to the substance, as defining the context) would not have significantly altered the observed data item. The contexts under which the analyst chooses to conduct the

hazard identification are developed using Working Table 1. Completion of Working Table 2 is discussed next.

4.2. Assessing Data Quality by Observational Context

The data available for each observational context need to be evaluated and summarized for application to hazard identification. If there is more than one data source on an item (i.e., data from more than one case) then the assessment of the data item should represent the coherence (or incoherence) of the composite information available. Discussion specific to summarizing across cases within a context will not be included here. Aside, perhaps, from application of meta-analysis for summarizing statistical outcomes across cases within a context, the ideas employed are similar to those for summarizing across contexts that will be discussed. Cases, however, are more homogeneous than contexts, which simplifies the judgments necessary.

Data characteristics to be assessed, shown along the top of Working Table 2, include (1) "completeness" (all data of the item were found), (2) "utility" (data are of high quality, applicable to hazard identification in the context observed, and statistically interpretable), (3) the observed effect (there is statistical evidence of a relationship between exposure and the data item), and (4) causality of the observed effect (the nature of the association between exposure and the observed effect is best described as...). These four characteristics of data quality as summarizing a data item will be discussed separately.

WORKING TABLE 2. DATA CHARACTERISTICS FOR OBSERVATIONAL CONTEXTS
CONTEXT NO. ___

Data Category/Item	Completeness (Hi/Me/Lo/No)	Utility[1] (Hi/Me/Lo/No)	Observed Effect[2] (Hi/Me/Lo/No)			Causality[3] (AA/CC/EC/OC)
			Exposure Effect	Context-Specific Measurement	Organism-Specific Measurement	
Tumor Response TR1 TR2 . . .						
Biophysical Effects BE1 BE2 . . .						
Pharmacodynamics PD1 PD2 . . .						
Host Characteristics HC1 HC2 . . .						
Related Substances Assessments RSA1 RSA2 . . .						

(continued on following page)

Working Table 2 (continued)

Data Category/Item	Completeness (Hi/Me/Lo/No)	Utility[1] (Hi/Me/Lo/No)	Observed Effect[2] (Hi/Me/Lo/No)			Causality[3] (AA/CC/EC/OC)
			Exposure Effect	Context-Specific Measurement	Organism-Specific Measurement	
Structure Activity Relationships SAR1 SAR2 . . .						

1 May be subdivided into additional categories when useful, e.g. "validity", "reliability", and "accuracy" may be judged separately for data items from laboratory sources.

2 Refers to effect on data item of exposure to agent. Footnotes with explanatory comments may be needed. When statistical measures are available, they may be more informative than a simple indication of Hi/Me/Lo/No, e.g., estimates or tests of statistical significance.

3 See Section 4.2.4 for explanation of the following choices available.

AA: Accidental Association
CC: Common Cause
EC: Empirical Causality
OC: Operational Causality

4.2.1. *Completeness*

A term commonly employed in discussions of both law and rationality is the principle of the "hard look" (see the appendix). The term is both descriptive and normative (i.e. implying that the "hard look" is necessary for a proper claim to rationality), introducing the intellectual obligation that rational discourse proceed from a base of observation and conceptual frameworks that is as large as possible without sacrificing the overall end of completing an analysis. The hard look may be taken to address three primary questions with respect to hazard identification. These are:

(1) Have all published studies (within a data item) been collected?
(2) Are published studies an accurate representation of all conducted studies (raising the issue of possible publication bias)?
(3) Do the collected studies represent accurately the phenomenon on which they supposedly report?

In the present document, the first two questions are taken to be questions of the completeness of data collection. Rationality then is weakened if the collected literature do not constitute an adequate sample of available studies. The third question is referred to here as an issue of utility. In this stage of the analysis, principles of study design and statistical sampling (a quantitative source of uncertainty through variability) are examined to determine whether the collected studies were capable of yielding an accurate representation of the empirical features of the data item being considered.

The data items included in Working Table 2 should include those reporting some outcome that may be a consequence of exposure to the agent of interest in the context of interest, e.g., from data categories such as tumor response, biophysical effects, concurrent environmental conditions, and pharmacodynamics (as applicable), as well as data items that are not context-specific but may have some bearing on a judgment of carcinogenicity from that context, e.g., data on structure activity relationships, host characteristics, and related substances

assessments if applicable to the context. In Working Table 2, the generic notation for data items (TR1, TR2, BE1, etc.) is used because the applicable data items will differ between contexts. For example, under "tumor response" the notation "TR1, TR2, ..." is replaced by specific data items from Table 2, such as "tumor incidence" and "age-at-appearance". The analyst then determines whether published studies have been adequately collected and published studies are an adequate representation of the full body of studies performed. For example, the study of ETS would be weakened by failure to include all of the published human epidemiologic data, or by demonstration that positive or negative studies had been consistently excluded from publication for some reason. In making this judgment the analyst will find it useful to reflect on several possible warrants for a claim to completeness:

(1) All studies (both published and unpublished) are known to have been collected.

(2) It is known that not all studies have been collected, but the collected studies represent a properly constituted random sample of available studies.

(3) All published studies have been collected and there is no publication bias.

(4) It is not known whether all studies have been collected, but it is judged that further searching would detract from the ability to carry out other tasks required in the analysis (i.e. contextual rationality, as discussed in the appendix).

Completeness is judged on a scale of "high", "medium", "low" or "no", depending upon the strength of the above warrants. An assignment of "high" implies that both published and unpublished studies have been collected, or that published studies have been collected and there is no publication bias. An assignment of "no" implies that there are no data available even after having adequately examined the literature. Having assigned a judgment of completeness to each data item, the analyst also assigns a judgment of completeness (on a scale of Hi/Med/Lo/No) to the data category.

4.2.2. *Utility*

Utility of a data item refers primarily to the quality of the process(es) under which observational data were generated, e.g., by an epidemiologic study, long term animal study, short term test, etc., *independent of the observed outcome* or of the completeness of the data category/item. The utility of a body of data increases with increasing ability of the studies to accurately and precisely define the underlying empirical properties under consideration in collection of the data item. Qualitative characteristics of interest are those related to methods and materials, including analysis and appropriateness of the conclusions drawn. For example, if observational information on the data item is from an epidemiologic study then the quality of the data item is judged *vis-a-vis* adherence of the study to principles of good epidemiologic practice. Similarly, if the data are from a long term animal study or a short term test, or are generated by some other means (as, for example, from consideration of structure activity relationships), the methods used to generate and analyze the data item determine its quality. The higher the quality, the more assurance that the observational information on the data item and the reported conclusions are a valid depiction and summary of what was claimed to be observed, aside from statistical sources of uncertainty or variability. Statistical uncertainty can further reduce the utility of a data item, quality not withstanding. Statistical error decreases with sample size, although size is not the only determinant. The smaller the sample size the lower the power (i.e., the less likely) to detect an association between the substance of interest and the data item. To summarize, utility should increase (1) as quality of the study (or other data-generating activity) increases, because the uncertainty due to methods and materials declines, and (2) sample size increases, because uncertainty due to statistical error declines.

Care needs to be taken to consider factors affecting utility that may not be foreseen or not readily apparent in the discussion above, such as applicability to the context in which the

data are being used in the analysis. For example, an epidemiologic study may be judged to be of high quality under one set of circumstances when evaluated with respect to study design, conduct, analysis, and conclusions, in the sense that all the recommended procedures for those circumstances are exercised with care and rigor. The quality may suffer, however, from the limitations of methodology under some other set of circumstances of interest to the analyst. Epidemiologic studies often have more than one objective and data for multiple endpoints may be collected simultaneously from a single questionnaire. Investigation of the substance of interest for hazard identification may be among the secondary objectives of a study or the available data may actually be a secondary data set from a previous study. The utility of the data items for hazard identification in such a case may be compromised even though the utility for the original end of the study was strong.

To illustrate, consider an epidemiologic study reporting tumor response from exposure to environmental tobacco smoke (ETS) as part of a larger objective to investigate health risks from indoor air pollution in general, of which the health effects from ETS is only one endpoint. If the study was conducted in a region with unusually high indoor air pollution from related substances, e.g., smoke from inadequate ventilation from indoor combustion of wood, coal, or other materials used in cooking or heating, then the study may have low sensitivity to detect an effect from ETS even though the quality for studying the effect of air pollution in general may be high with respect to study design, conduct, analysis, etc. and the sample size may be extremely large. What needs to be recognized in this case is the presence of factors that methodology may not adequately address. A "noisy" background may make detection of an effect from ETS very unlikely, contrary to the appearance of adequate sample size based on power calculations. The application of statistical methods to adjust for confounding, and hence to "correct" for the presence of other pollutants that may produce detrimental health effects

similar to ETS, may be inadequate in such a situation. The difficulty does not lie in the accuracy of the statistical methods, but in application to situations (contexts) that exceed their limitations (i.e., the assumptions on which the validity of the methods are based are violated sufficiently to produce misleading results). Consequently, an epidemiologic study might be judged high in quality against the textbook criteria for good epidemiologic practice and be of adequate sample size, but warrant a lower judgment for utility (i.e., usefulness, applicability, value, contribution to the weight of evidence, etc.) within a given context defined by Working Table 1 and employed in Working Table 2. Technically, this illustration is not a counterexample to assessing quality, as discussed above. It does indicate, however, that care needs to be exercised in making judgments of utility for data items, and this applies in varying degrees to all the data items of Working Table 2.

The next two sections sketch some basic features of animal studies and epidemiologic studies to consider in assessing utility. The discussion is largely summarized from the OSTP guidelines for chemical carcinogens (Fed. Reg., 1985, p.10372 - 10442), and interspersed with the authors' opinions or material from other sources. There is considerable literature on both topics, particularly related to the design and conduct of epidemiologic studies, to which the reader is referred. The OSTP guidelines include numerous references that may be useful. Some more recent sources on the design and conduct of animal studies are described in the following section. Material on the design and conduct of epidemiologic studies is readily available in textbooks, journals, and other sources.

4.2.2.1. Utility of Animal Studies Long-term animal exposure studies often provide much of the data available for hazard identification. There are numerous references that discuss principles of design and protocol for animal studies that may be useful in guiding a judgment of quality

(utility) for data from animal studies (Code of Federal Regulations 1983a, 1983b; Hamm, 1985; Prejean, 1985; Boorman et al., 1985; U.S.DHHS, 1984; Krewski et al., 1990; Haseman, 1985; Portier and Hoel, 1983; IARC, 1986). Critical design criteria are summarized below under five of the six headings used by the OSTP guidelines (Fed. Reg., 3/14/85, p.10411-), and much of the discussion is excerpted from that source. Only the OSTP guideline related to the selection of species and strain of animal is omitted here since that factor is related to inter-context extrapolation (see Chapter 5) rather than utility as developed here.

(1) Animal Care and Diet. Housing, purity of diet, water and air, proper housing and care of animals, control of intercurrent diseases or parasites, and controlled exposure to the test agent are critical to valid testing. Care must be taken to avoid bias in selection and allocation of animals between controls and treated groups, or in placement of cages.

(2) Test and Control Groups. It has been recommended that each dose group and a concurrent control contain at least 50 animals so that enough animals are available at the end of the study for pathological examination. Fewer animals reduces statistical power to detect an effect and reduces the expected number still alive at the end of the study, reducing the number exposed for "life". With interim sacrifice of animals, the number should be larger. The control group and treatment groups should be treated identically aside from exposure to the agent for valid comparisons.

(3) Dose Levels, Frequency and Route of Exposure. Two or three dose levels in addition to the concurrent control group are typically used. The highest dose currently recommended is the maximum tolerated dose (MTD), a dose just high enough to elicit

minimal toxicity without significantly altering animal survival rates under lifetime exposure. Lower dose levels are typically at 1/2 or at 1/3 and 2/3 of the MTD. The rationale for dose selection should be clearly stated. The rationale for use of the MTD is to maximize the chance of detecting a carcinogenic effect. Lower doses provide information on the shape of the dose response curve, for dose-response assessment, but are also important for hazard identification. Multiple doses are needed to implement tests for trend and to indicate nonlinear behavior that may influence extrapolation of results across species or to low exposure levels. The route of administration used for the bioassay ideally corresponds to the route of exposure of humans. However, so long as the route of administration used on test animals results in absorption, distribution, and metabolic activation (if required) of the substance, the test results are generally regarded as relevant for hazard identification in humans.

(4) Duration of study. Test animals are normally exposed for most of their expected lifespan, which means a study duration of about two years for rats and mice. Treatment is typically begun soon after weaning. Termination of the study is ordinarily acceptable when the number of survivors in the low dose or control group is reduced to 20-25%. If high-dose toxicity affects survival in that exposure group, the study is still continued with the remaining animals at lower exposure levels and in the control group. A negative test is ordinarily accepted by regulatory agencies if: (1) no more than 10% of any group is lost due to autolysis, cannibalism, or management problems; and (2) survival of all groups (per sex per dose) is no less than 50% at 80 weeks for hamsters, 96 weeks for mice, and 104 weeks for rats.

(5) Data Collection and Reporting. Animals need to be observed on a regular basis for indications of toxic effects or disease, and for autolysis or cannibalism. Body weights and food intake should be recorded on a regular schedule. Clinical signs and mortality should be recorded for all animals, with special attention paid to visible signs of tumor development.

4.2.2.2. Utility of Epidemiologic Studies. The obvious advantage of epidemiologic data or animal bioassay data is that observations are on human subjects, relieving the uncertainty of extrapolation of results across species. On the other hand, the researcher has no control over the exposure environment as in animal studies (completely removing the claim of operational causality discussed in Section 4.2.4.) and often cannot completely specify "Concurrent Environmental Conditions" (see Table 2). Exposure to the substance of interest may vary between study subjects, and temporally for any given subject, and environmental conditions and other factors of potential importance are not constant (see Section 5.4.4.). Those limitations not withstanding, however, it is often difficult to obtain accurate and reliable data on the factors of interest. Consequently, some types of epidemiologic studies are more useful for hazard identification than others. For example, descriptive studies may utilize the correlational (or ecological) approach, in which the rate of disease in a population is compared with the spatial or temporal distribution of suspected risk factors. This type of study is helpful in developing or refining hypotheses for further investigation, but is not very useful for hazard identification since causality is unresolved. Data are collected on whole populations instead of individuals, providing only correlational evidence that is too broad-based for inference of a causal association between increased incidence of cancer (or other effect) in one population and increased exposure to the substance of interest (again, see Section 4.2.4.).

Analytical epidemiologic studies are the principle means for determining the human health hazards of specific environmental agents. In contrast to a descriptive survey, data are obtained on disease occurrence and putative risk factors for specific individuals, using mainly the case-control or cohort method. Case-control studies start by identifying persons with a particular disease (cases) and a group of similar persons without the disease (controls). Information on past exposure to the agent of interest and other potential risk factors is collected from which statistical comparisons are made between cases and controls. If the frequency of exposure to the agent is higher in cases than in controls, after adjustment for other risk factors that might produce the disease (for hazard identification, the disease is a cancer endpoint, such as lung cancer in the studies of passive smoking, although it could be more specific, such as lung adenocarcinoma), then the analyst must consider whether the observed outcome indicates an association between the agent and the disease. The case-control approach is well suited to studying relatively rare diseases, where either (1) exposure to the agent is common, as in studies of environmental tobacco smoke or menopausal estrogens, or (2) exposure is rare but accounts for a large portion of a particular cancer, such as liver angiosarcoma in the study of vinyl chloride.

By contrast, cohort studies start by identifying a group of individuals with a particular exposure and a similar group of unexposed persons, followed by examination of both groups over time to determine subsequent health outcomes. The incidence of disease in the exposed and unexposed groups are then compared. These investigations may be based on current exposure and future health outcomes (prospective cohort study), but more commonly they utilize past exposure information and disease occurrence (retrospective cohort study). Instead of an unexposed comparison group, general population mortality or incidence rates (specific for age, sex, race, and calendar time) are often used to determine the expected number of cases of

diseases. Cohort studies are expensive, complex, require large numbers of exposed persons, and long periods of follow-up. They are less subject to sources of bias (to be discussed further), however, the risks attributed to a particular exposure can be estimated directly, multiple health outcomes (including multiple cancer endpoints) can be assessed, and temporal relationships such as latency period and duration of effect can be evaluated. The intervention study is a third type of analytical epidemiologic study, especially useful in confirming causal relationships suggested by case control or cohort studies (see particularly the discussion of operational causality in Section 4.2.4.). This approach may be applied in programs designed, for example, to reduce cigarette smoking.

The qualitative integrity of the design, conduct and methods of analysis in epidemiologic studies is an essential determinant of the utility of the study results for hazard identification. The capability to eliminate sources of bias and confounding as possible explanatory factors in lieu of exposure to the agent being studied may depend heavily on characteristics related to study quality. In both case-control and cohort studies, confounding cannot normally be removed by appropriate study design alone, i.e., possible confounders need to be controlled in the analysis, as possible. Since data are required, the collection of data on possible confounders enhances the utility of a study (see "Concurrent Environmental Conditions" in Table 2 and the discussion of causality in Section 4.2.4.). Bias is more of a concern in case-control studies than cohort studies, arising largely from the case-control design. In particular, care should be taken to avoid bias (to the extent possible) in the selection of cases and controls for study and in the collection of data on exposure and related risk factors. In general, qualitative characteristics of an epidemiologic study involve design, execution, and analysis. Methodologic criteria for evaluation of quality are readily available in textbooks and other references on epidemiology. Two IARC scientific publications specific to the application of case-control and cohort studies to

cancer research that may be particularly useful are by Breslow and Day (1980, 1987), although they focus on statistical methods. Additional sources that discuss the application of epidemiology to cancer research include Cook (1982), Peto(1985), Hoel (1985), Day (1988), Wald and Doll (1985), NRC(1986), Muir and Higginson (1985), Hoel and Landrigan (1987), Mettlin (1987), Cone et al. (1987), Krewski et al. (1990), Morris (1990), and Sorsa et al. (1990).

Table 4 contains a checklist of items for methodological critique of studies as constructed by Spitzer et al. (1990). Not shown in the table is the choice of responses to each item (yes, uncertain/incomplete/substandard, no, don't know/not reported, N/A, N/C, and space for comments). A worksheet for review of case-control studies of ETS was developed by Spitzer et al (1990) around five categories, including: (1) General. Study objective, primary or secondary data set, meaning of terms ("nonsmoker", "exposed to ETS", etc.), recall span (duration since ETS exposure), type of exposure (cigarette, pipe, etc.), classification of ETS exposure in unmarried women. (2) Data Collection. Inclusion/exclusion criteria for cases (and separately for controls), source of subjects (hospitals, community, etc.), incident cases (Y/N), control sampling (cumulative/density/matched/unmatched), method of collection (face-to-face, telephone, self-administered questionnaire, medical records, etc.), verification of data with other sources (Y/N), sample size, attrition in selection and follow-up, source of data on subject (subject or proxy), exposure periods (adulthood, childhood, etc.), exposure sources (smoking by parents, spouse, household members, etc.), age. (3) Clinical data. Method of verification of primary lung cancer, airway location of lung cancer (central or peripheral), prevalence by tumor type (squamous cell, adenocarcinoma, etc.). (4) Statistical Analysis. Raw data, unadjusted (crude) statistical analysis (type and outcome of statistical tests for effect, for trend, etc.), adjusted statistical analysis (same tests but adjusted for confounding factors). (5) Dependent variables used in matching (of controls to cases), in analysis, and used or discussed otherwise. A checklist (such as Table 4) or

TABLE 4. CHECKLIST FOR METHODOLOGICAL CRITIQUE OF EPIDEMIOLOGIC STUDY

1. Random assignment, properly done
2. Suitable choice of reference group
3. Similar methods of data collection for all groups
4. Proper sampling or suitable assembly of comparison group
5. Sample size
   a. enables adequately precise estimates of priority variables found to be significant
   b. enables adequate precision in secondary variables reported (confounding variables or incidental findings)
   c. power reported for nonsignificant findings
   d. power declared *a priori*
   e. clinical or practical significance of statistically significant differences set forth or justified
6. Criteria for definition or measurement of the outcomes are objective or verifiable
7. Definition of exposure; unambiguous and measurable
8. Measurement of exposure; accurate and verifiable
9. Blind assessment
10. Observation bias minimized by design or accounted for in analysis
11. Selection bias accounted for
12. Objective criteria for eligibility of subjects (inclusion and exclusion)
13. Attrition rates (%)
   a. response rate
   b. losses to follow-up
   c. other
14. Known confounders accounted for
   a. by design
   b. by analysis
15. Any methods to attempt comparability between groups, other than randomization
16. Comparability of groups under comparison demonstrated
17. Appropriate statistical analytic plan
   a. evidence that *a priori* hypotheses being tested
   b. correct method used
   c. adjustment made for
      - multiple comparisons
      - simultaneous multiple range testing
   d. display of raw data permits assessment of actual measures and adjustments or transformations made
18. Conclusions supported by data presented
19. Reproducibility of method(s)
20. Generalizability of results
   a. from sample(s) to parent population
   b. from sample(s) to any relevant population
21. Other, specify

Source: Spitzer et al. (1990)

worksheet for items that may be important to judging utility of a study facilitates inclusion of significant items and detection of items not addressed in a study that should be.

The utility of an epidemiologic study is also affected by its statistical power to detect an effect of dose on the health endpoint of interest (this, and the following comment, also applies to the utility of animal studies). If statistical significance is not achieved (typically a p-value < 0.05 is interpreted as "significant"), then it is of interest to know the power of the test to detect the minimal effect level that would be considered consequential. It is not uncommon for an estimate of relative risk to be "high", e.g., 2 or larger, but not significant because of inadequate power. One cannot simply conclude, however, that if the sample size had been sufficiently large, then the estimate would remain unchanged but become statistically significant. One could argue that statistical power only affects the utility of a study when the outcome is nonsignificant. Since the outcome is not known in advance, however, factors affecting statistical power must be taken into account during the design phase. Factors affecting power include: (1) Size of the study group and control group(s). (2) Variability of the background cancer rate. (3) Criteria for significance (i.e., predetermined p-value to reject the hypothesis of no effect, technically called the "test size"). (4) Magnitude of the expected association between exposure and effect. (5) Design of the study and the statistical techniques used for analysis. (Marsh and Caplan, 1986, as reproduced in Morris, 1990, p.271).

4.2.2.3. Summarizing Utility For the data item under consideration, the utility of the set of available data is judged on a scale of "high", "medium", "low" or "no". An assignment of "high" implies the data are characterized by an appropriate study within the prespecified context, including both qualitative features (use of proper experimental and/or epidemiologic techniques) and quantitative features (e.g. power). As either of these characteristics weaken,

the assignment of utility is lowered due to decrease in the ability of the data to accurately correspond to the property being measured. It is important to note here that the assignment of utility does not include the issue of confidence intervals on point estimates, which are discussed in the next section on observed effect.

4.2.3. *Observed Effect*

Observed effect for a data item indicates the level of empirical support for some form of association between the data item and the agent of interest, *independent of the judgment of utility and completeness for the item*. As discussed above, when the context for Working Table 2 includes more than one case a composite result for the data item in that context needs to be formulated by the analyst for each data item. For example, suppose that for epidemiologic studies of lung cancer and ETS, each country for which there are studies is defined as a context and the intra-country studies are the cases for the context. Under data category "tumor response" TR1 in Working Table 2 might denote "relative risk". A composite estimate, confidence interval, and significance level (p-value) might be formulated for relative risk by pooling the statistical results (*not* the data) on relative risk from the component studies. Data items under the category "biophysical effects" would be treated similarly, although statistical results from each study may be limited to tests of significance. Since estimates, confidence intervals, and tests of significance contain different information, the analyst may choose to enter these pooled statistical outcomes for a data item in place of a more qualitative choice from (Hi/Me/Lo/No), or perhaps present the statistical summary in a footnote. Since the end of hazard identification is to produce a claim that the substance does (or does not) induce a change in cancer incidence, rather than an estimate of the magnitude of any change, loss or the

quantitative information in summarizing the strength of association is not significant (unless magnitude of change is deemed a measure of the existence of an association).

Only compatible statistical tests may be pooled, e.g., the outcome from a trend test cannot be combined with the outcome of a pairwise comparison of cases and controls. Technical deliberation may be required to form an overall judgment of observed effect. Data items under "tumor response" may apply to an epidemiologic study or to an animal study, depending on the context. Statistical methods, of course, differ accordingly. Several of the references cited in the discussion on "utility" above (Sections 4.2.3.1 and 4.2.3.2) include material on statistical methods and interpretation of results as well as aspects of design that may be helpful. References for long term animal studies include Gart et al. (1979), Mantel (1980), Peto et al.(1980), Clayson et al. (1983), Haseman (1983, 1984), Haseman et al. (1984), Park and Kociba (1985), Maronpot(1985), Huff et al. (1986), Farrar and Crump (1988), Krewski et al. (1988), Archer and Ryan (1989), Bickis and Krewski (1989), Haseman et al. (1989), Haseman (1989), Krewski et al. (1989), Portier and Bailer (1989), Haseman (1990), and Wahrendorf (1990). For epidemiologic studies, the two IARC volumes by Breslow and Day (1980 & 1987) are particularly oriented toward cancer studies, although there is a vast array of readily available literature. Data on biophysical effects can generally be statistically analyzed with the effect of interest replacing cancer as an endpoint, although categorical data may be more common since observations may include severity rankings by a pathologist, e.g., classification hyperplasia as severe, moderate, mild, or absent.

Data for relating the strength of association of a data item with exposure level are less likely to be available (or particularly meaningful) for the six remaining data categories (host characteristics, pharmacodynamics, concurrent environment conditions, related substance assessments, and structure activity relationships). In particular, the data categories "related

substances assessments" and "structure activity relationships" would not include data items for which "exposure effect" has meaning for the factor under consideration in the analysis (although it will have meaning for the "related substance") itself. Host characteristics may predispose some populations to higher exposure effects than typical, which needs to be taken into consideration (see examples in Tables 5 and 6). Consequently, there will be several forms of "observed effect" summaries possible, as indicated in Working Table 2.

The first form of summary statement arises from the "tumor response" and "bioeffects" data for which exposure-response observations are available. In this case, the desired judgment is whether an association has been noted and, if so, the strength of that association. A judgment of "No" indicates that there is no association, which this does not mean the data item will not affect assessment of support for claims of carcinogenicity. A judgment of "Hi", "Med", or "Lo", on the other hand, indicates some association (from "strong" to "weak") between exposure to the agent and presence of the observed factor (tumor incidence, cellular transformation, etc.). This form of statement is given in the column "Exposure-Effect" under "Observed Effect".

The second form of summary statement arises from other data categories in which the observed factor (data item) has been measured in the *presence of the agent of the level of exposure characterizing the context*. For example, the data item "bioactivation enzymes" may contain data on the presence or absence of necessary metabolic enzymes under conditions of exposure. This form of statement is given in the column "Context-Specific Measurement" under "Observed Effect".

The third form of summary statement arises from data similar to those in the second form with the exception that they (the former data) have not been obtained under conditions of exposure. An example might be the presence of repair enzymes for DNA in unexposed organisms (the organisms specifying the context). The analyst must already have judged, of

TABLE 5. DEVELOPMENTAL PROCESSES THAT ENHANCE SUSCEPTIBILITY TO ENVIRONMENTAL POLLUTANTS

High-Risk Groups	Estimated Number of Individuals in United States Affected	Pollutant(s) to Which High-Risk Group is (may be) Increased Risk
	Developmental Processes	
Immature Enzyme Detoxification Systems	Embryos, fetuses, and neonates to the age of approximately 2-3 months	Pesticides, polychlorinated biphenyls (PCBs)
Immature Immune System	Infants and children do not reach adult levels of IgA until the age of 10-12	Respiratory irritants
Deficient Immune System as a Function of Age	Progressive degeneration after adolescence	Carcinogens, respiratory irritants
Differential Absorption of Pollutants as a Function of Age	Infants and young children	Barium, lead, radium
Retention of Pollutants as a Function of Age	Individuals above the age of 50	Fluoride, heavy metals
Pregnancy	Approximately several million females per year in the U.S.	Anticholinesterase insecticides, carbon monoxide, lead
Circadian Rhythms Including Phase Shifts	All people have certain periods of the day when they are more susceptible to challenge	Hydrocarbon carcinogens and probably most other pollutants
Infant Stomach Acidity	Infants	Nitrates

Source: Calabrese, E.J. (1984)

TABLE 6. PREVELANCE OF SUBGROUPS
HYPERSUSCEPTIBLE TO EFFECTS OF COMMON POLLUTANTS

Hypersusceptible	Prevalence	Chemicals
Embryo, Fetus, Neonate	Pregnant Women: 21/1000	Carcinogens, solvents, CO, mercury, lead, PCBs, pesticides
Young Children	Ages 1-4: 70/1000	Hepatotoxins, PCBs, metals
Lung Disease		Ozone, Cd, particulates, SO_2, NO_2
Coronary Heart Disease	Coronary Heart Disease: 16-27/1000	Chlorinated solvents, fluorcarbons
Liver Disease	Liver Condition: 20/1000	Carbon tetrachloride, PCBs, insecticides, carcinogens

Source: U.S. EPA (1984)

course, that such data items are applicable within the exposure conditions characterizing the context. If the observed property is likely to be strongly affected by exposure to the agent, this initial judgment of inclusion in the context would have been invalid. This third form of summary statement is given in the column "Organism-Specific Measurement" under "Observed Effect".

For both the second and third forms of summary statement, the judgment is not one of an association. It is, instead, a measurement of some property related to theories of the etiology of cancer. The judgment is that the observed property is "strongly present" (Hi), "moderately present" (Med), "weakly present" (Lo) or "not present" (No) for the substance within the context under consideration. In each case, the "strength" of the "presence" is related to the effect on a claim of carcinogenicity. For example, substances that are strongly absorbed by the G.I. tract, tenaciously retained in the target organ, and readily activated to the biologically significant metabolite through biotransformation in the organism would be assigned a value of "Hi" in each of these three "observed effects" columns. Conversely, substances that are not deposited readily in the lung following inhalation, that are not neutrophilic, or do not partition from water to oil would be assigned a value of "Lo" or "No" in each of these three "observed effects" columns. For reviews of the role of the various data items in judgments of carcinogenicity, the reader should consult Andersen (1981), Andersen (1989), Anderson et al. (1980), Barrett and Wiseman (1987), Clayson (1987), Clewell and Andersen (1989), Gehring et al. (1978), Gerlowski and Jain (1983), Gillette (1976, 1984), Hattis (1990), Keck (1981), NRC (1986, 1987), US EPA (1989) and US FDA (1982).

4.2.4. *Causality*

This section reviews the separate judgments to be made in warranting the claim that an association noted in a given data item is or is not an indication of causality (again, see Working

Table 2 and the "Exposure-Response" data items under the column "Observed Effect"). These causal judgments take two primary forms. The first refers to properties of the particular study from which the observation was obtained. This component focuses only on characteristics of the study (data item) itself indicative of causality pertinent to the current discussion. The second refers to expectations formed independent of the particular study, and is more properly contained within the issue of theory-based inference raised in Section 5.2.4. These two forms of causal warrant are discussed together here only due to their conceptual link. It should be borne in mind that only the first mode of warranting causality is to be used in forming the judgment of causality for Working Table 2.

The study results themselves may provide warrants for a judgment of causality. The intent of the warrants is to distinguish between the following cases (proceeding from the weakest to the strongest warrant):

(1) Accidental association. This is a judgment of non-causality. It is addressed through reflection on statistical properties of the study and is summarized in statistical measures of the significance of the association (see the discussion in Section 4.2.2.). Such a judgment is warranted when the statistical properties are insufficient to rule out the null hypothesis.

(2) Common cause. This is a judgment that the factor under consideration (such as exposure to the substance of interest) does not cause the effect noted in the study, but the factor of interest is caused by the same process that produces the true causal factor. For example, particulates in ETS have been taken as a causal factor in lung carcinogenesis (U.S. EPA, 1990). This claim to causality will be weakened if both particulates and gases have a common cause in the burning of cigarettes, and it is the gas which is the true carcinogen. Still, the use of particulate measurements as an indicator of hazard in an environment will be appropriate if it is premised that the factor under consideration (particulates) is in common cause with the true factor (gas). To be in common cause implies that the factor under consideration is present simultaneously with the true causal factor, since whatever produces the former also produces the latter. Mitigation strategies based on lowering of the former factor will produce the desired effect since any mitigation of the former implies mitigation of the latter factor. The utility of common cause arguments is greatly weakened, however, when it is possible that the factor under consideration might in some cases exist independently of the true causal factor. Common cause arguments, based on strength of association only, is a form of epistemic instrumentalism as discussed

in the appendix. A judgment of common cause is warranted when the factor of interest is accompanied by other factors, as observed in the data category "Concurrent Environmental Conditions", and where existence of the factor of interest is caused by some process that also gives rise to one or more of the "concurrent" factors.

(3) Empirical causality. In this instance, features of the study are used to warrant the claim that the associations noted are physically causal rather than accidental or the result of common cause. These features are given significance through theories of carcinogenesis describing properties of study findings that are to be expected in instances of a causal connection. The primary indications of causality then are (1) appearance or disappearance of the effect when exposure is historically present or absent (respectively), (2) the temporal pattern of appearance of the effect (i.e. the latency period or the hazard function), (3) appearance of the tumors at the site of the BSDR produced by the exposure, and (4) there are no "Concurrent Environmental Conditions" judged capable of explaining the observed change in prevalence.

(4) Operational causality. This is the strongest warrant for a judgment of causality. The evidence invoked results from a proper experimental setting in which the researcher has produced deliberate and intentional manipulations (operations) of the factor of interest to yield the effect. The physicality of the manipulation is essential in the warrant, and the strength of warrant increases as the analyst increases confidence that only the intended manipulation occurred.

The four categories of causality above are to be warranted through reference to empirical properties of the studies. For example, criteria for causal inference in epidemiologic studies have been developed (Hill, 1965), qualified (Evans, 1978; Weiss, 1981) and discussed in a context of inductive and deductive logic (Maclure, 1985). These empirical criteria, used in assigning causality for Working Table 2, are not discussed here further due to their extensive review in the existing literature. A causal connection may, however, also be supported through arguments based on theory, often referred to as biological plausibility (in fact, this form of causal argument must be excluded from consideration in generating Working Table 2, since it is used in subsequent parts of the analysis discussed later). The reasoning here is not that the study characteristics support the contention of causality (although this is not excluded), but rather that prior understanding of the observed phenomenon leads to an expectation of a causal

connection. Etiologic theories, and observational items from categories other than the category under consideration, then might be invoked as further support for the claim to a causal connection in the "tumor response" and "biophysical effects" data. For example, the causal connection between exposure to ETS and lung cancer (as given in the data category "Tumor Response") may be warranted partially by reference to observations on active smoking (in the "Related Substances" category) and theories founded on the concept of cigarette equivalence (Thorslund, 1990). This prior expectation of a causal connection must be assessed, of course, both for the exposure factor of interest and for any other factors invoked as competing explanations of the observed effect. It will be useful (as discussed in Chapter 5) to think of the empirical and theory-based warrants as providing a system of warrants within which the conceptual coherence of the claim to causality may be judged. In any event, theory-based warrants for causality (i.e. biological plausibility arguments) constitute an instance of Theory-Based Inference towards the claim of carcinogenicity, and discussion of this mode of warranting is deferred until Section 5.2.4.

5. WARRANTING CLAIMS TO CARCINOGENICITY

The preceding two chapters have dealt with the first two of the four tasks for hazard identification specified in Section 1.3., specifically the development of a taxonomy of claims of carcinogenicity and the assemblage and assessment of observational evidence. The taxonomy does not need to be reconsidered with each new agent evaluated for carcinogenicity, although it could be altered as beneficial. Establishment and assessment of the informational base is, of course, specific to the agent and/or context of interest. Working Tables 1 and 2 suggest formats to assist with the organization and evaluation of observational evidence suitable for input to the last two tasks concerned with warranting claims of carcinogenicity. The first of these two

remaining tasks (i.e. Task 3) is the warranting of carcinogenicity claims within each observational context, i.e., for each context in Working Table 1 for which there is observational evidence. The completed Working Table 2 for each such context is utilized with the relevance strategies to be described in this section to form weight-of-evidence judgments of the support for claims of carcinogenicity shown in the taxonomy (Table 1). This third task is specific to intra-context warranting of claims, i.e., the claims apply to an observational context listed in Working Table 1. For example, suppose the basis for Context 1, as described in Working Table 1, is a long-term animal study with Wistar rats exposed to controlled levels of the agent. All of the relevant context-specific data items from that study would be included in Working Table 2 (for Context 1), as well as data items from other sources that are not context-specific but considered directly applicable, e.g., historical control rates of tumors in Wistar rats, chemical properties of the exposure agent, pharmacodynamic information, etc. (as summarized in the "Organism-Specific Measurement" column under "Observed Effect"). The objective of Task 3 of hazard identification is to assess support for carcinogenicity of the substance of interest *vis-a-vis* the species, exposure levels, and other relevant conditions specific to each context, without extrapolation of evidence across contextual boundaries. Working Table 3 (to be described) provides a format to facilitate Task 3 for each observational context.

These working tables (given by one Working Table 3 for each observational and target context) serve as input, along with information from other sources that may be considered applicable (described in Section 5.4.), to the fourth and final task of hazard identification concerned with inter-context warranting of claims. Observational evidence is often unavailable or inadequate for the target context of exposure to humans, or humans at low exposure levels typical of environmental conditions. In that case support for claims of carcinogenicity within the target context depends on the intra-context judgments formed in Task 3 for observational

WORKING TABLE 3. INTRA-CONTEXT SUPPORT FOR CLAIMS OF CARCINOGENICITY[1]
CONTEXT NO. ___

Reference Strategy	I.O.[3]	Claims of Carcinogenicity[2]								
		Increases Incidence of Cancer	Classification(s)				Stage(s)		Mechanism(s)	
			Complete	Partial	Mixer	Helper	Neo. Conv.	Neo. Devel.	Geno-toxic	Non-genotoxic
Direct Empirical (D.E.)										
Semi-Empirical Extrapolation (S.E.E.)										
Empirical Correlation (E.C.)										
Theory-based Inference (T.B.I.)										
Existential Insight (E.I.)										
Column Summary										
Overall Summary										

1 Top half of each entry is completed using dose-response data observed in the context. Bottom half of each entry includes "floater-data" as well (see text).

2 Choices for cell entries are Hi/Me/Lo/No. See text of Section 5.1. This assignment is made independently of the assignment of intellectual obligation.

3 "Intellectual Obligation" (Hi/Me/Lo/No). See text of Section 5.1. This assignment is made independently of the assignment in the cells of the table (footnote 2).

contexts, and the availability of supporting data needed to extrapolate claims of carcinogenicity across contexts. The third and fourth tasks are described in detail in Sections 5.2., 5.3. and 5.4. In the interim, the discussion turns to a general consideration of strategies by which observational data are given relevance to tasks. Throughout the present chapter, a common example of the carcinogenicity of inhaled radon is employed, with the observational base and warrants being those discussed in the NRC (1988) and NRC (1991) reports on radon. The former report is concerned primarily with a discussion of data in the categories of "tumor response", "related substances" and "biophysical effects", and the latter report with a discussion of the role of "pharmacodynamics", "host characteristics" and "concurrent environmental conditions" data.

5.1. Relevance Strategies

Tasks 3 and 4 of hazard identification require the analyst to judge the levels of support for claims of carcinogenicity. Fundamental to this warranting of claims, whether intra-context or inter-context, is the epistemic basis on which observational information is judged to be supportive evidence of a claim. The five alternative categories of evidential reasoning available to the analyst, referred to here as relevance strategies, are described below. The characteristics of the available observational data used in a warrant for a claim of carcinogenicity, as assessed by context in Working Table 2, in conjunction with a judgment of the support for any background premises necessary for use of the data in lines of reasoning, determine the *strength* of the warrant; the relevance strategy employed is the *type* of warrant. More than one type of warrant may be applied to data to support a claim of carcinogenicity and, as discussed in a later section (Section 5.5.), the coherence of support across warrants is an important consideration in the overall judgment of the epistemic status of a claim.

(1) Direct Empirical (D.E.). This category involves the claim that a given observation constitutes a direct observation of the property under debate (such as carcinogenicity). For example, observational data on inorganic arsenic includes epidemiologic studies relating skin cancer to arsenic levels in well water and clinical observations linking treatment of acne with an arsenic solution to increased occurrence of skin cancer. Interpretation of these studies alone as evidence for a claim that arsenic increases the risk of skin cancer in humans (at some exposure level) would be an example of warranting by direct empiricism. Similarly, there exist studies relating exposure of human populations to radon in the home and elevations in lung cancer incidence (see, for example, the review by Samet, 1989), although the coherence of these studies is low since they do not uniformly indicate an increase in lung cancer incidence within exposed groups. The strength of the warrant would be based on the characteristics of the study data. For example, the strength of a warrant would be increased by establishing that the studies were of high quality and the elevated incidence was not the result of confounding factors (such as the presence of other chemicals that may play a role) or unusual antecedent conditions (such as atypical host characteristics). It must also be demonstrated that the studies were conducted within the defining context. Clearly, as these background premises are better satisfied, the strength of the direct empirical warrant improves. In the case of the NRC (1988) report, direct empirical studies are mentioned but their utility is judged to be too weak to justify use of the direct empirical warrant.

(2) Semi-Empirical Extrapolation (S.E.E.). This category involves the claim that a given observation was not obtained under the desired context of exposure, but that any result of the observation may be extrapolated to the desired context. Basically, the claim is to having observed a pattern in the available exposure-response data which may be followed to the described context. For example, dose-response curves obtained on mining populations and experimental animals exposed to high concentrations of radon have shown no evidence of a threshold down to exposures an order of magnitude above those occurring in homes (NRC, 1988). While not providing direct empirical evidence of effects in the home, the data do provide a less convincing warrant for such effects, conditional on the acceptance of the background premise that the mechanism of carcinogenicity is fundamentally similar at high and low exposures and that the exposure-response pattern is evident in the available data. Clearly, as any patterns in the data become better established, and as the background premises already discussed in (1) above are supported, the strength of the semi-empirical warrant improves. Examination of Figures 2.A.1. through 2.A.4. in the NRC (1988) report suggests that exposure-response patterns are not clear in the data but generally are monotonic without an evident threshold. A more germane finding is that trend tests are statistically significantly positive across the four studies examined.

(3) Empirical Correlations (E.C.). This category involves the claim that a particular type of observation (such as of in-vitro transformation) is correlated with the desired property (such as carcinogenicity). There is no claim to understanding why the correlation exists, this claim being the function of the fourth category of relevance strategy discussed below. Clearly, as the strength and specificity of the

correlation improves, so does the strength of the warrant for empirical correlations. It is important to note that observational evidence on substances other than the one of interest generally is used in warranting claims to correlations (see the "Related Substances" data category in Working Table 2), requiring a judgment as to how sets of substances are to be chosen in developing correlations. The NRC (1988) report does not explicitly cite empirical correlations as a warrant, although (as discussed in Section 5.2.3.) correlations between induction of chromosomal aberrations and carcinogenicity are cited to strengthen the claim that radon is expected to be a carcinogen or at least induces a BSDR.

(4) Theory-based Inference (T.B.I.). This category involves the use of data (such as in-vitro cellular transformation or bio-activation processes) which do not yield claims to observing cancer but do represent claims to observing an effect related to the etiology of cancer. In this case, the role of background premises appearing in theories of carcinogenesis becomes of central importance. Clearly, the strength of theory-based inference improves as the evidence for the background premises embodied in theories improves. It is important to note here that other bodies of observational evidence, perhaps drawn from substances other than the one of interest, might be invoked as warrants for a particular background premise appearing in an etiologic (mechanistic) theory. For example, the premise concerning a role of adduct formation in neoplastic conversion might be warranted by data from other substances in which adduct formation has been shown to play such a role. In the NRC (1988) report, information on cellular radiobiology, particularly as this relates to the demonstrated ability of radon to induce chromosomal damage and transformation of cells, is invoked to support the claim that radon is a carcinogen. The invocation of theory-based warrants justifies a claim to both empirical and conceptual success (see the appendix).

(5) Existential Insight (E.I.). As discussed in the appendix, there may be times when a scientific expert judges that an observation suggests carcinogenicity, but the expert is unable to explain how this judgment arose. The judgment is a product of personal experience and a-logical (as opposed to illogical) reflection, as in the case of "engineering judgment". While not strictly a logical warrant, the analyst still may deem the judgment rational and wish to factor it into the analysis. Clearly, the strength of this "warrant" improves as it is better shown that the expert possesses (1) the necessary prior experience from which such judgments might spring and (2) the skill of reflecting on experience and producing reliable existential judgments (see the discussion in Section 5.2.). The only explicit use of this mode of warrant within the NRC (1988) report is in the assignment of priors for the Bayesian analysis of the mining data.

The preceding discussion of relevance leaves two issues unaddressed. First, some measure of the quality of a given relevance strategy (for each of the five strategies above) must be assigned. This measure, again on a scale of low to high, is determined through reflection on

the quality of the observational data employed in using the strategy (from Working Table 2) and by the strength of evidence for the background premises required by invocation of the strategy (see the discussion of premises in Section 5.2.). The necessary background assumptions are specific, in turn, to the particular taxonomic carcinogenicity claim, but are of the general form noted in the discussions of the five relevance strategies above. For example, use of information on in-vitro production of chromosomal aberrations by radon (Brandom et al., 1978) in warranting a claim of carcinogenicity requires such background premises as: (1) the necessary chromosomal damage occurs both in-vitro and in-vivo, (2) chromosomal aberrations result in necessary DNA alterations, (3) the DNA alterations induce transitions, etc. An assignment of "low" implies that the background premises have not been verified for the substance of interest and the desired taxonomic claim, and an assignment of "high" implies strong verification of these premises for the particular claim. These judgments are made in isolation from consideration of the strength of the claim that chromosomal aberrations have, in fact, been observed.

The second issue of warranting the use of relevance strategies (as opposed to the strength of those strategies irrespective of use) relates to the existence of five different strategies for relevance. Regardless of the strength of warrant for the premises underlying a given relevance strategy (such as the outcome of examining direct empirical evidence or existential insight), the analyst must provide an indication of the degree to which the analyst deems it rational (a priori) to base decisions on a particular mode of warranting claims to carcinogenicity. Another way of putting this is that the analyst determines the degree to which availability of a given strategy is an "intellectual obligation" in rationally justifying a claim to carcinogenicity. Again, the scale is from low to high. An assignment of "low" implies that the analyst does not find a particular relevance strategy to be reliable generally and/or a proper conception of the "intellectual obligations" (Alston, 1985) required of a rational decision-maker (as might be the

case in existential insight). An assignment of "high" implies that the analyst finds a particular relevance strategy to be reliable generally and/or a proper conception of the intellectual obligations required of a rational decision-maker (as might be the case in direct empirical evidence). It is important to bear in mind that this assignment is independent of the evidence available for the premises in each particular strategy in each particular case (i.e. for each substance considered). It is formed, instead, on the basis of the epistemic status of a relevance strategy in general as discussed in the appendix. It is a form of "intellectual obligation" imposed consistently on the analyst during any act of warranting and requires coherent application of the same obligation to each specific instance of warranting. In other words, broad principles of epistemic reasoning should be applied consistently throughout the analysis in the absence of counter-arguments as to why a relevance strategy might be more acceptable in one act of reasoning than in another. This assignment is included here due to the fact that some individuals may generally trust or distrust particular forms of reasoning (such as reliance on existential insight or etiologic theories) and yet may apply this trust or distrust inconsistently within an analysis or across analyses. While not discussed explicitly in the NRC (1988) report, the degree of intellectual obligation implicitly assigned to "Direct Empirical" warrants must have been low given the fact that the entire analysis of risk is based on semi-empirical evidence (i.e. the evidence from high exposures in the mines) and theory-based inference (particularly through use of the radiobiological experimental data).

5.2. Application of Relevance Strategies to Warrant Intra-Context Claims of Carcinogenicity

The role of data categories in carcinogenicity claims is described in Table 7, with reference to the relevance strategies defined in the last section. The relevance strategies available to each combination of data category and carcinogenicity claim are displayed in Table

8 for reference. In this section the relevance strategies are applied to the data items in Working Table 2 to formulate judgments of the level of support under each type of evidential reasoning for each claim of carcinogenicity in the taxonomy. Working Table 3 is suggested as a format for displaying the judgments. Cells in the bottom row of the table, labeled "overall assessment", are for summary judgments from each column after the main part of the table is completed and assessed for coherence across the five potential types of relevance strategy (see discussion in Section 5.5.).

The reader will note that the cells of Working Table 3 (aside from the "summary" cells) are divided by a dashed line. This division arises due to the presence in Working Table 2 of some data specific to an exposure-context and others specific to the organism but not measured in the context of exposure at the level of interest (contained under the sub-column "Organism-Specific Measurement" in Working Table 2). The latter data may appear in several contexts (i.e. several instances of Working Table 2), allowing the same data to play multiple roles in the analysis perhaps out of proportion to the significance of the data (since they do not represent empirical values obtained under exposure conditions). To prevent this, the judgments for cells in Working Table 3 first are formed on the basis of only data from columns "Exposure-Effect" and "Context-Specific Measurement" in Working Table 2. These judgments, placed into the cells of Working Table 3 above the dashed lines, "root" the claims of the analyst in the primary data. The secondary data, given by the "Organism-Specific Measurement" column in Working Table 2, then are analyzed to "perturb" the primary judgments of the analyst, strengthening or weakening the pre-existing assignments. These modified assignments then are placed into the cells of Working Table 3 below the dashed lines. It is these modified assignments that will be used in producing the summary assignments for the table (see Section 5.5.). The assignments themselves are on a scale from "Hi" to "Lo" as discussed in Section 5.1.

TABLE 7. RULE OF DATA CATEGORIES IN CARCINOGENICITY CLAIMS

Data Category[1]	Role of Data Categories in Carcinogenicity Claims
Tumor Response[2]	Direct empirical support for "increases cancer", conditional upon causal premise. Direct empirical support for "complete carcinogen" if concurrent environmental conditions or previous/later initiation/promotion conditions were not required to yield cancers. Direct empirical support for "partial carcinogen" if previous/later initiation/promotion conditions were required to yield cancers or if observed effect is only an increase in multiplicity of tumors. Direct empirical support for either "mixer" or "helper" claims if concurrent exposure must be present (but incapable of distinguishing between the two claims). Provides empirical correlation and/or existential insight warrant for any of the above claims. This data category has no bearing on claims to "stages" and "mechanisms".
Biophysical Effects	Data items provide Theory-Based Inference warrants for the claims of "increases incidence with cancer", "complete carcinogen" and "partial carcinogen". The ability to distinguish between a "complete" and "partial" carcinogen requires Biophysical Effects data indicating that the factor induces all necessary transitions or only a subset of such transitions. The data items can also provide empirical correlation and existential insight warrants for the above claims, but direct empirical and semi-empirical warrants are not available. If the data are an indicator of transitions, they may provide direct empirical, theory-based inference, empirical correlation and/or existential insight warrants for the "stage" claims. If the data items are an indicator of conditions necessary for transitions, they may provide the above warrants for the "mechanism" claims. If any of the data items were obtained under conditions of concurrent exposures, they may provide direct empirical, empirical correlation and/or existential insight warrants for the classification as "mixer" or "helper".
Host Characteristics	This data category does not provide a warrant for claims to carcinogenicity. The data affect only the decision to include a given organism/study within a given context.
Pharmaco-dynamics	These data support only the contention that a BSDR is produced in the organism. Therefore, they do not directly warrant intra-context claims to carcinogenicity. They can, however, provide a theory-based warrant for the contention that any effects observed ("tumor response" and/or "biophysical effects") are causally connected to exposure by warranting the contention that at least a BSDR is produced in the organism.
Concurrent Environmental Conditions	These data support the contention that an observed effect (e.g. cancer prevalence) was connected causally to exposure to the factor of interest. The data also provide Theory-Based Inference support for the "classification" of carcinogenicity claims if it is shown that concurrent exposures are (mixer/helper) or are not (complete/partial) required to yield the observed effect. The data have no implications for the other claims to carcinogenicity.

(continued on following page)

Table 7. (continued)

Data Category[1]	Role of Data Categories in Carcinogenicity Claims
Structure Activity Relationships	These data provide Theory-Based Inference, Empirical Correlation, and/or Existential Insight warrants for the claims of "increases incidence of cancer", "complete carcinogen" or "partial carcinogen". The specific claim warranted depends upon the structural feature of the molecule and the presence of the appropriate mechanism of action within the organism. If, as is often the case, the structure indicates an initiating agent, the data provide support for a partial carcinogen with no ability to distinguish further between a partial and complete carcinogen. The data do not pertain to the remaining classifications (mixer/helper). They do, however, provide warrants for the claim that a substance acts on a given stage or by a given mechanism, particularly with respect of initiation (neoplastic conversion and genotoxicity).

1 See Working Table 2.

2 Quantitative measures are shown under "data items" for "tumor response" only. Quantitative measures commonly used are not given in other data categories, only the endpoints of interest.

TABLE 8. RELEVANCE STRATEGIES AVAILABLE FOR CLAIMS OF CARCINOGENICITY

Data Category	Claims of Carcinogenicity[1]								
	Increases Incidence of Cancer	Classification(s)				Stage(s)		Mechanism(s)	
		Complete	Partial	Mixer	Helper	Neo. Conv.	Neo. Devel.	Geno-toxic	Non-genotoxic
Tumor Response	D.E. E.C. E.I.	D.E. E.C. E.I.	D.E. E.C. E.I.	D.E. E.C. E.I.	D.E. E.C. E.I.	D.E. E.C. E.I.	D.E. E.C. E.I.		
Biophysical Effects	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	D.E. E.C. E.I.	D.E. E.C. E.I.	D.E. T.B.I. E.C. E.I.	D.E. T.B.I. E.C. E.I.	D.E. T.B.I. E.C. E.I.	D.E. T.B.I. E.C. E.I.
Host Characteristics	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Pharmaco-dynamics	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Concurrent Environment Conditions	N.A.	N.A.	N.A.	T.B.I.	T.B.I.	N.A.	N.A.	N.A.	N.A.
Related Substances Assessments	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.
Structure Activity Relationships	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	N.A.	N.A.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.

[1] N.A.: None Available
D.E.: Direct Empirical
S.E.E.: Semi-empirical Extrapolations
E.C.: Empirical Correlations
T.B.I.: Theory-based Inference
E.I.: Existential Insight

5.2.1. *Direct Empirical Warranting*

Direct empirical warrants are available at times when a data item may be taken as directly indicative of carcinogenicity (i.e. as a warrant, in and of itself and with minimal theoretical interpretation, for a carcinogenicity claim). As shown in Working Table 3, the separate claims to be warranted are that:

(1) The substance (or exposure factor) increases the incidence and/or time of appearance of cancer in a specific context (exposure level and species).

(2) The substance is a partial carcinogen, complete carcinogen, "mixer" or "helper".

(3) The substance acts on the organism through neoplastic conversion and/or neoplastic development.

(4) The substance acts on the organism through genotoxic and/or non-genotoxic mechanisms.

Turning first to the claim that a substance increases the incidence of cancer, the direct empirical warrant requires observations on the incidence of cancer within the context of interest. As a result, only observations in the category of tumor response may provide the basis for a direct empirical warrant. The strength of the warrant then increases as (1) the completeness of the observational base in a data item is improved, (2) the quality (utility) of the observational studies increases, (3) the strength of the association (dose-effect) increases and (4) the claim to a causal connection between exposure and the observed incidence is supported (see Working Table 2). In the case of radon, direct empirical evidence was available (Samet, 1989), but the utility, strength of association, causal connection and coherence was low to moderate.

For the distinction between the various classifications of carcinogens (complete, partial, mixer and/or helper), direct empirical warrants require additional background premises and, hence, additional observational evidence. For classification as a complete carcinogen (in the row labeled "Direct Empirical" in Working Table 3), the necessary premise is that the substance may increase cancer without the presence of concurrent exposures within the considered context.

This requires tumor response data in which (1) the role of the substance as the sole causal agent (aside from contributions from non-experimental factors) is well established, warranted by the data category "concurrent environmental conditions" as summarized in the causality claim in Working Table 2 and (2) cancer incidence was increased, warranted by the "tumor response" data category. Classification as a partial carcinogen requires introduction of the premise that deliberate (experimental) exposure to either an initiating, promoting or progressing agent must accompany exposure to the substance of interest, requiring assays demonstrating empirically that exposure to the substance must be accompanied by (previously, concurrently or subsequently) exposure to either an initiating or promoting agent. Radon has been suggested to act primarily through initiation at low doses and, hence, is to be considered primarily a partial carcinogen. Classification as a mixer requires introduction of the premise that there be concurrent exposure to another substance that is not itself a carcinogen, and empirical evidence that exposure to the substance of interest must be accompanied by "concurrent environmental conditions" (see the discussion of causality for Working Table 2) constituted by exposure to that other non-carcinogenic substance. For classification as a helper, the same premises apply with the exception that the "other" factor must be a carcinogen.

For the distinction between the stages (neoplastic conversion and/or neoplastic development), the analyst must introduce the additional premise that the substance acts explicitly to yield neoplastic conversion (as in the case of radon) and/or neoplastic development. Direct empirical support for these claims may be taken as observation of biophysical effects (see Working Table 2) such as cellular transformation (NRC, 1988; Woodruff, 1990b; Barrett et al., 1986a, 1986b; Bartsch and Malaveille, 1990) and progression of preneoplastic lesions (Bannasch et al., 1987), or from characteristics of the tumor response such as alterations in the time-to-

appearance but not the incidence of tumors (indicating an effect on the "speed" of neoplastic development) etc., following exposure within the desired context.

Finally, the mechanistic distinctions require premises that the substance acts through changes in DNA (genotoxicity, as in the case of radon (NRC, 1988)) or changes in other biological structures/functions (non-genotoxicity) (Barrett and Wiseman, 1987; Barrett, 1987; Butterworth, 1990). Direct empirical support for premises is, again, given by observations of biophysical effects, although of a more fundamental nature than those required in distinguishing stages. The observational support for genotoxicity generally is taken as formation of DNA adducts, DNA breakage, chromosomal aberrations, oncogene activation/deactivation and/or base pair alterations (for reviews of the evidence for causal roles, see Barrett and Wiseman, 1987; Barrett, 1987; Butterworth, 1990; Belinsky et al., 1987c; Farber, 1987; Lawley, 1987; Morris, 1990; Perera, 1987). Evidence for the role of chromosomal damage is radiation-induced cancer is summarized in Hall and Freyer (1991). The observational support for non-genotoxic mechanisms generally is taken as changes in inter-cellular communication, interference with hormonal control, immune system destruction, hyperplasia induction, etc. (for reviews of evidence for causal roles, see Langenbach et al., 1988; Lewis and Adams, 1987; Loury et al., 1987; Moolgavkar, 1986; Perera, 1984; Scribner et al., 1987; Slaga, 1984; Swenberg and Short, 1987; Trosko and Chang, 1988). Crawford-Brown and Hofmann (1990) have proposed a non-genotoxic mechanism by which cytotoxicity of radon results in promotional effects.

5.2.2. *Semi-Empirical Warranting*

Since the present section (5.2.) is directed towards intra-context claims, instances of semi-empirical extrapolation are not appropriate at the second level of warrant. This form of warrant will be described in Section 5.3., which deals with extrapolation across contexts

(including extrapolation across exposure levels within another homogenous context, the subject of semi-empirical warrants).

5.2.3. *Warranting from Empirical Correlations*

Correlational warrants require knowledge aside from the evidence on tumor response for the agent of interest. Construction of sets within which correlation coefficients (strength and specifity) between a data item and a carcinogenicity claim may be constructed requires premises detailing how those sets should be constructed. In other words, it must first be specified how a data item is to be chosen as a candidate for inclusion in the set of items for which the correlation has been developed. These premises, in turn, are related to theories of carcinogenesis. As a result, empirical correlations may be closely related to etiologic theories. Although distinctly different in nature, it is convenient to discuss these two categories (theory and correlation) together because empirical correlations have been used to support or question cancer paradigms and thus have had some influence on the direction of their development. For example, only a few years ago many analysts would have appealed to theory-based inference to warrant a claim of "non-carcinogen" if an agent tested negative in short term tests (STTs) for mutagenicity, such as the Ames salmonella bioassay. In this instance, the major premise of the theory was that carcinogens acted through mechanisms of geonotoxicity to alter DNA structure/function. But these predictions did not correlate well with observational evidence from long-term animal studies, causing existing paradigms of cancer mechanisms to be modified or refined. By the same token, theories of carcinogenesis (such as the influence of mitogenesis) have altered the construction of sets within which correlation coefficients are developed. This example will be discussed further, but it illustrates the inter-relationship of the predictive capacity of empirical correlations and theory-based inference following from the partial

warranting of cancer paradigms by empirical evidence, i.e., when theory-based predictions are not consistent with the empirical record, then the theory is re-examined (and vice versa).

The introduction of bacterial and cell culture techniques to measure chemically-induced mutations led to a widely held belief that mutagenicity and carcinogenicity were highly correlated. In particular, the Salmonella/mammalian microsome test of Ames et al. (1975), took center-stage as a predictor of carcinogenicity based on the short-term test for mutagenicity. Somatic cell mutation was often considered a necessary initiating step in cancer. In a recent review, Ashby (1990) summarizes "the simple paradigms of 1975" as follows: "Human carcinogens are also carcinogenic to rodent; rodent carcinogens can be assumed to be human carcinogens. The majority, if not all, organic rodent carcinogens are reactive to chromosomes or their constituent DNA (genotoxic), usually following appropriate metabolic conversion. The Salmonella mutation assay, coupled with a few other *in vitro* genotoxicity assays, is sufficient for the detection of such genotoxic chemicals. With cancer, it is unwise to assume a safe-dose level of human exposure. Thus, at its simplest level the activity of a chemical in the Salmonella assay implies a cancer hazard to exposed humans." In contrast, in what he refers to as the "worst-case scenario in 1990" , Ashby notes that "No new human carcinogens have been defined since 1975: many new rodent carcinogens are now known, many of which current data indicate to have no cancer hazard for humans...a large and growing number of rodent carcinogens are non-genotoxic,. i.e., they cannot be detected using current genotoxicity assays. Further, a large and growing number of *in vitro* genotoxins appear to be devoid of rodent carcinogenicity."

The loss of simplicity in the last fifteen years is partially explainable by the increased awareness that cancer does not develop by a single mechanism, common to all species and routes of exposure wherein an agent may increase cancer incidence. Indeed, the primary biological activity of an agent or a metabolite may not be genotoxic, as indicated by point

mutations, insertions, deletions, or changes in chromosome structure or number. Chemicals that exhibit such genotoxic activity can usually be detected by assays that measure reactivity with DNA, induction of mutations or DNA repair, or cytogenetic effects (Butterworth, 1990). The earlier conceptual simplicity of a high correlation between mutagenicity and carcinogenicity was further discredited by evaluation of the results from long-term rodent assays conducted by NCI and NIEHS. In an initial article results of four widely used *in vitro* assays for genetic toxicity applied to 73 chemicals recently tested in the two-year NCI/NIEHS long term rodent studies were compared (Tennant et al., 1987a, 1987b). The chemicals selected were well characterized in rodents for carcinogenicity or non-carcinogenicity. Of the four assays (for induction of mutations in Salmonella (SAL) and mouse lymphoma L5178Y cells (MLA), and induction of sister chromatid exchanges (SCE) and chromosome aberrations (ABS) in Chinese hamster ovary cells). It was found that SAL detected only about half of the carcinogens as mutagens and the other three assays (MLA, SCE, and ABS) did not complement SAL, i.e., there was no gain from using other any of the other three assays in addition to SAL. These results were confirmed by examination of 41 additional chemicals (Zeiger et al., 1990).

The disappointing sensitivity of genetic toxicity tests to predict rodent carcinogenicity was compounded by the results for specificity reported earlier by Shelby and Stasiewicz (1984) who found that more than 60% of 80 rodent non-carcinogens had been found active in at least one of the four *in vitro* tests. These results suggest that *in vivo* testing may be complementary to *in vitro* tests. NTP is currently evaluating *in vivo* tests for chromosome aberrations and micronuclei as a means of confirming *in vitro* mutagenicity (Zeiger et al., 1990). It may also be noted here that the second Collaborative Study on the Assessment and Validation of Short-Term Tests for Genotoxicity and Carcinogenicity (CSSTT/2) concludes that "*in vivo* tests have a vital role to play in hazard assessment. This role is to define which chemicals, identified as genotoxic from

in vitro tests, are active *in vivo* and, thus, are those most likely to present a carcinogenic/mutagenic hazard to mammals, including humans" (WHO, 1990).

Further correlational analyses of NCI/NTP rodent studies have been conducted that are informative for claims of relevance in hazard identification based on empirical correlations and/or theory-based inference (Ashby and Tennant, 1988; Ashby, 1990; Tennant and Ashby, 1991; and Ashby and Tennant, 1991). Results reported in the last citation, entitled "Definitive relationships among chemical structure, carcinogenicity, and mutagenicity for 301 chemicals tested by the U.S. NTP", are excerpted in the following discussion. A high correlation was observed between structural alerts to DNA reactivity and mutagenicity, but neither predicted rodent carcinogenicity effectively. It appears that certain tissues are only sensitive to genotoxic carcinogens, e.g., zymbal's gland and the lung. Other tissues in which cancers were observed are subject to both genotoxic and non-genotoxic mechanisms. It is noted that "non-genotoxic rodent carcinogens cannot be automatically neglected as possible human carcinogens. However, if non-genotoxic carcinogens induce their effects via an interaction between the test agent and the sensitive rodent tissue, the relevance of that interaction to human tissues should be studied, particularly at low dose-levels, before a human carcinogenic hazard is assumed. Such questions of extrapolation are distinct from those posed by genotoxic rodent carcinogens where comparative metabolism and the intrinsic genotoxic potency of the test agent are the critical parameters, DNA itself being common to rodents and man."

"Non-genotoxic" is a non-descript classification simply referring to any mechanism other than genotoxic and, as discussed above, understanding the specific mechanism may be key to predicting whether a rodent carcinogen is a cancer hazard to humans or whether increased cancer incidence at a high exposure level constitutes a hazard at environmental concentrations (a particular concern with respect to mitogenic action). As noted by Butterworth (1990) and

other authors, a mechanism may appear to include both genotoxic and nongenotoxic elements. For example, non-DNA reactive promoters that induce cell proliferation can yield mutagenic events and chromosomal alterations secondary to that proliferation while mutagens given at cytotoxic doses may induce cell proliferation. In practice, however, it may be useful to apply a weight-of-evidence approach to classify a chemical as genotoxic or nongenotoxic, noting that the latter category is a class of mechanisms that may not depend solely on chemical properties, but the context of exposure (concentration, species, route of administration, etc.). Butterworth (1990) suggests that such an approach might include "examination of constituents of the molecule for prediction of DNA reactivity, overall results from cell culture assays, activity in the whole animal, dose and target organ specificity, histopathological evaluation of tumor development, and other obvious potential mechanisms of carcinogenicity (such as cell proliferation) (Ashby and Tennant, 1988)". Cohen and Ellwein (1990) also recommend classifying chemical carcinogens as genotoxic or nongenotoxic, with further division of the latter category by mechanisms of action, if known. In particular, it is of interest to note whether the chemical acts through a cellular receptor or a non-cellular mechanism. Cohen and Ellwein note that agents acting through specific receptors tend to be active at low doses, making it unclear whether there is a threshold level of exposure.

Ashby and Morrod (1991) summarize some current concerns on carcinogenicity testing and their interpretation for hazard identification in light of recent developments discussed above, principally the limited predictive capacity of short term tests for genetic toxicity and heightened awareness of the prevalence and diversity of nongenetic mechanisms in long term animal studies. Their suggested approach to testing for carcinogenicity has implications for the type of data needed to warrant claims of carcinogenicity, of interest in the present report. Ashby and Morod recommend an initial step of inspecting the chemical structure of an agent for

sites of DNA-reactivity, followed by assessment of genotoxicity using *in vitro* tests and if necessary, short-term rodent tests. This will identify potential genotoxic carcinogens. An issue in genetic toxicology is the use of *in vivo* tests to separate *in vitro* genotoxins that are rodent carcinogens from those that are noncarcinogenic to rodents, and the use of complementary genotoxicity assays to detect the few Salmonella-negative genotoxins, such as benzene. The authors note that the vast majority of classical human and rodent chemical carcinogens, together with most of the NTP two-species carcinogens, are overtly genotoxic *in vitro* and *in vivo*. The next step is to evaluate the remaining genotoxins for a range of toxicities associated with non-genotoxic rodent carcinogens, noting that agents found to be inactive are probably not carcinogenic. The authors provide a list of 14 potential non-genotoxic indicators, i.e., there is some supportive evidence of association for the indicators but none has been established definitively as a predictor. Without an explanatory mechanistic link with cancer at present, it is suggested that the absence of each indicator is supportive of "probable non-carcinogenicity". It is noted that many issues about non-genotoxic carcinogenesis remain to be understood. (Note: The authors claim nothing new in their suggested scheme. Citations and credits have not been included in the abbreviated discussion provided here.)

The discussion above is intended to provide an introduction to some of the issues that need to be addressed in hazard identification and to give some idea of the type of mechanistic information that may be usefully applied in an weight-of-evidence approach to hazard identification. Attention is now directed to additional sources of empirical correlations for warranting claims of carcinogenicity.

The 114 NCI/NTP studies used for comparison of four *in vitro* tests of genetic toxicity by Zeiger et al. (1990) and Tennant et al. (1987a), described above, are evaluated for correlational characteristics in Haseman and Clark (1990). The 114 chemicals are all those tested by

NCI/NTP during a specified time period, thus eliminating the possibility of selection bias. The NCI/NTP claims of carcinogenicity are as follows: 67 (59%) were carcinogens in at least one of the four sex-species groups (male/female, rats/mice), the evidence for 17 (15%) was found to be equivocal, and 30 (26%) were judged to be non-carcinogens. For interpretation of these figures and the ones to follow, it is necessary to understand how NTP assigns a "category of evidence" to a chemical based on evidence in rodents studies (usually of rats and mice of both sexes). As described by Haseman and Clark, the final decision as to whether a study was positive, equivocal, negative, or inadequate is a matter of scientific judgment. Biological factors are considered as well as the statistical outcome, including: whether a dose-response was observed; whether pre-neoplastic lesions were observed; the historical control rate, i.e., the accumulated evidence from previous control groups regarding the spontaneous occurrence of observed neoplastic lesions; biological characteristics of the lesions observed; the survival of dosed and controlled animals; whether tumor latency was affected by dose; the multiplicity of site-specific neoplasia; the adequacy of the experimental design and conduct of the study; and the consistency of occurrence across sexes within a species, and across all four sex-species groups. Consequently, "rodent carcinogenicity" refers to an assigned category based on human evaluation of quantitative and qualitative characteristics of observational data from the long-term animal studies. With awareness of factors considered in forming a claim of rodent carcinogenicity, the correlations noted by Haseman and Clark in their review of the evidence from 114 chemical studies will be described.

The inter-species concordance is 69% (compared to 74% for the whole NCI/NTP database reported by Haseman and Huff (1987)). The concordance of 69% is comparable to the concordance of 66% between the Salmonella-assay and rodent carcinogenicity for these chemicals reported in Zeiger et al. (1990). The 67 carcinogens are unlikely to represent an

equal risk to humans, according to Haseman and Clark, who conclude that carcinogens producing effects at multiple sites and/or in multiple sex-species groups are probably more important from a public health standpoint than single-sex and single-site carcinogens, or those where carcinogenic sites were accompanied by organ toxicity. Haseman and Clark draw some conclusions regarding the effect of chemicals that are toxic, but this does not refer to evidence of organ toxicity. They refer to toxic potency of the chemical, i.e., chemicals with low maximum tolerated doses (MTDs) are more "toxic" than those with high MTDs. The intended significance is not clear, but it is reported that most (67-88%) of chemicals positive in STTs (referring to the four *in vitro* tests described above that were applied to this set of chemicals, namely SAL, SCE, ABS, and MLA) were also "toxic" and the majority (62-78%) of negative STT chemicals were "nontoxic". Overall concordance between toxicity and rodent carcinogenicity was 65%, about the same as the same as reported for STT outcomes and carcinogenicity in this set of chemicals. It was found that all four STTs were correlated with chemical toxicity, particularly SCE and MLA ($P < 0.001$). The Salmonella-assay is more predictive of carcinogenicity among the toxic chemicals (65% positive SAL in carcinogens vs. 12% in noncarcinogens) than in non-toxic chemicals (with corresponding percentages of 13 of 3). ABS demonstrated a similar pattern but with less predictive capacity, while SCE and MLA results showed no evidence of association with rodent carcinogenicity even within classes of toxic or non-toxic chemicals.

In another review of the NTP database, Hoel et al. (1988) address the issue of whether organ toxicity, particularly at high doses, may be associated with judgments of rodent carcinogenicity. Of principle concern is the extent to which a secondary mechanistic process, such as cytotoxicity with resultant compensatory cell proliferation, may be responsible for tumor induction in chemicals classified as rodent carcinogens. Since such a mechanism would be unlikely to occur at environmental exposure concentrations, the classification of carcinogenicity

is more nearly an artefact of exposure to high doses than an indicator of a potential cancer risk to humans. A total of 378 two-year studies on 99 chemicals conducted by the NTP were studied, 53 (54%) of which were considered to be carcinogenic in one or more of the laboratory-animal studies. The authors conclude that "only seven of the 53 positive carcinogenicity studies exhibited the types of target organ toxicity that could have been the cause of all observed carcinogenic effects. Furthermore, no apparent difference in mutagenicity as measured by the Ames Salmonella assay was observed between 'high dose only' carcinogens and the entire set of carcinogens." Some qualifications are contained within the article, however, relating to issues that "certainly merit further investigation in order to make a more definitive statement about toxicity and carcinogenicity". The reference stresses the importance of clearly defining what is meant by "organ toxicity" and "preneoplastic lesion" in discussions of either.

Wilbourn et al. (1986) considered the validity of extrapolating results from long-term carcinogenicity tests in animals to humans by reviewing the 41 IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans (7-47) that had been published or were in press. The IARC claims of carcinogenicity are the judgment of working groups of experts who consider both epidemiologic and animal data. It was found that "84% of the 44 exposures with sufficient or limited evidence of carcinogenicity to humans also have some carcinogenic activity in animals. The remaining chemicals and complex mixtures have not been adequately tested for carcinogenicity in animals. In no case was there 'no evidence of carcinogenicity' ". It is also found that "for many exposures causally related to human cancer, there is a target organ in common between humans and at least one animal species, despite many inherent physiological differences."

Allen et al. (1988) conducted quantitative comparisons of carcinogenic potency in animals and humans for 23 chemicals for which suitable animal and human data exist. These

comparisons, based on the TD_{25}, were found to be strongly correlated. The best prediction of human results from animal data was achieved by utilizing data from several routes of exposure and employing the assumption that animals and humans are equally sensitive to a carcinogen when dose is measured in units of mg/kg body weight/day.

5.2.4. *Warranting by Theory-Based Inference*

While Section 5.2.3. contained material related to etiologic theories, the invocation of theories in that section was intended (1) as a tool for establishing sets within which empirical correlations between a data item and carcinogenicity could be defined and (2) as a tool for examining the role of existing correlations in testing axioms (premises) for theories. The current section discusses the role of theories in directly warranting claims of carcinogenicity through judgments of conceptual understanding (rather than strength of empirical correlation). This warrant requires explicit use of both observational evidence for effects believed to constitute part of the causal sequence leading to cancer (aside from tumor response) and etiologic theories of cancer uniting the evidence and giving it relevance to the task of supporting a claim of carcinogenicity. The two broad premises that must be introduced are (1) the substance induces a biologically significant dose-rate (BSDR) following exposure and (2) the resulting BSDR induces cancer in one of the taxonomic forms of claims of carcinogenicity (see Working Table 3). Both premises must, of course, be warranted within the desired context. For example, in the case of radon (NRC, 1988; Hall and Freyer, 1991; NRC, 1991), the observation of chromosomal aberrations, even at exposures below those for which direct empirical (epidemiologic) evidence was available, is taken as evidence that at least the necessary first steps in inducing cancer occur at these low exposures. This suggests that a BSDR is produced at

these exposures and that this BSDR results in damage believed to be etiologically significant in the production of transitions towards cancer.

It is important to note here that theory-based inference does not employ direct empirical evidence of carcinogenicity as contained in tumor response data (see Working Table 2), although it may employ direct empirical evidence in separately warranting the two broad premises above (as was done in the case of radon). The intent, instead, is to determine the implications of other available data (biophysical effects, pharmacodynamics, etc.) within the framework of etiologic theories. This isn't to say that the analyst ultimately ignores the tumor response data. Those data appear, instead, as direct empirical warrants for the claims. In addition, "Related Substances Assessments" data might provide observational warrants for the premises underlying etiologic theories themselves, conditional upon acceptance of the further premise that there is a common etiologic foundation to carcinogenicity in the contexts of the "related substance" and the substance of interest to the analyst (as when active smoking data are used to produce a theory of carcinogenicity for ETS, although it should be borne in mind that the premise of a common etiologic basis for the action of different carcinogens is increasingly being questioned in light of findings on the multiple mechanisms through which carcinogens may act). For radon, the general finding of the carcinogenicity of all forms of ionizing radiation has been invoked (NRC, 1988) as a warrant for the claim that radon (which yields ionizing radiation) is a carcinogen. It might be useful to think of theory-based inference as providing a prior expectation of carcinogenicity for a substance, prior to consideration of the direct empirical evidence. Theory-based inference also warrants a claim to conceptual understanding, allowing the analyst to assert not just that the substance has been observed to induce cancer (summarized in the "Direct Empirical cells of Working Table 3) but that the etiologic role of the substance is understood within the framework of existing theories. Such an argument is

particularly strong in the case of radiation, where mechanistic understanding is much better developed than for chemical carcinogens. The ability to assert scientific understanding is an important component of scientific rationality and was raised initially in Section 4.2.4. and in more detail in the appendix.

The first premise (that a BSDR is produced following exposure) is itself warranted by observations contained under pharmacodynamics in Working Table 2. These data are summarized for radon exposures in the latter NRC report (NRC, 1991). The role of each of these observations within the reasoning of the analyst (i.e. in supporting the judgment that a BSDR is produced) is displayed in Table 7 and is not repeated here. The warrant for the first premise may take any of the five forms discussed in Section 5.1. For example, the analyst may have observations of a BSDR being produced in the desired context, leading to a direct empirical warrant for the first premise (but not, of course, for the claim to carcinogenicity, since this discussion is contained within theory-based inference with respect to the carcinogenicity claims). The analyst may have observed only that the necessary enzymes for bioactivation are present, resulting in a theory-based warrant for the first premise or in a warrant of empirical correlation or existential insight. Again, any of the five warrants (relevance strategies) may be available for warranting the first premise. For radon (NRC, 1991), it has been experimentally confirmed that radon progeny are deposited in the lung and produce a dose-rate of radiation to sensitive cellular subpopulations. The form of warrant for the premise is not included formally in the working tables, but is factored into the judgment of the analyst concerning the strength of the warrant for "Theory-Based Inference" in Working Table 3.

An identical discussion applies to the second premise, i.e. that a BSDR (when present) is capable of inducing cancer in one of the senses noted in the claims to carcinogenicity found in Working Table 3. This second premise also may be warranted by any of the 5 relevance

strategies and, as in the first premise, the form of warrant for premise 2 is factored into the judgment of the strength of the "Theory-Based Inference" warrant of Working Table 3 (without appearing explicitly in that table). For radon (NRC, 1988), this premise is warranted primarily be experimental observations that irradiation of cells in-vitro yields transformed cells. The role played by each potential observational category (and, hence, data item) in any line of reasoning lending support to the second premise is displayed in Table 7.

Many of the effects potentially measured in a study of exposures to humans, animals or cell lines are not of cancer induction directly but rather of biophysical phenomena believed to play a causal role in carcinogenesis. Use of these data to support a claim of carcinogenicity within theory-based inference requires introduction of a premise that the bioeffect at least is an indicator of one or more of the transitions leading to a tumor. In other words, the etiologic role of the observed bioeffect within the process of carcinogenesis appears as an explicit premise associated with the etiologic theory used in drawing a "Theory-Based Inference" in the cells of Working Table 3. Support for this premise, in conjunction with observation of the bioeffect itself, then strengthens the second premise that the substance-induced BSDR yields transitions.

The role of the bioeffects data in supporting the premise that BSDR results in cancer occurs on two conceptual levels. The first level relates to the specific stages of carcinogenesis, taken here to be neoplastic conversion and development (see the discussion in Chapter 2). The primary data of use at this level are measurements of preneoplastic lesions, initiation ability, promotional ability, in-vitro cellular mutation and in-vitro cellular transformation, all of which are discussed in the NRC 1988 report on radon. Initiation assays are taken to support a contention that the substance of interest induces at least the conversion stage, requiring the additional premise that the further transition (development) occurs with some non-zero background rate or substance-induced rate. A similar comment applies to the results of

promotional assays. Preneoplastic lesions are taken as an indication of both conversion and promotion having taken place, requiring a premise that progression eventually would occur in some fraction of the preneoplastic lesions if sufficient follow up time was available (a particularly important premise in extrapolation from studies on short-lived animals to humans) (Bannasch et al., 1987)). It is important here to bear in mind the specificity of transitions induced by a given carcinogen, which might imply that a partial carcinogen can complete the necessary transitions only for specific forms of the other transitions not induced by that partial carcinogen. In other words, existing data on carcinogenicity indicate that it is not generally true that a given promoter (for example) is capable of promoting all initiated cells. It may, instead, be capable of promoting only specific forms of initiation-related damage. It also should be noted that bioeffects data support claims that a substance induces some of the necessary transitions, but not that it is a complete carcinogen (unless the bioeffect is taken to be the entirety of effect necessary for cancer, with only a probabilistic component remaining in the production of a frank tumor).

The second level at which bioeffects data may support the premise (that BSDR results in cancer) relates to observations on processes leading to transitions between stages, but not observations of the stages or transitions themselves. It typically is premised that neoplastic conversion results from changes to DNA (see Chapter 2) with a less well-established link between DNA changes and neoplastic development. As a result, the ability of a substance to produce DNA adducts, DNA breakage, DNA gene mutations, activation of oncogenes, or de-activation of repressor genes may be taken as evidence of neoplastic conversion if it is premised that either (1) these bioeffects are correlated with conversion or (2) these bioeffects play a causal role in conversion. The first premise is warranted by reference to correlational studies using principles analogous to those as discussed in Section 5.2.3. The second is warranted by

reference to (a) explicit mechanistic understanding of the process of conversion or (b) a claim to existential insight using principles analogous to those discussed in Section 5.2.5. Structural characteristics of molecules have also been correlated partially with the conversion activity of those molecules (Ashby and Tennant, 1988, 1991; Ashby et al., 1989). Again, the ability of radon-associated radiation to produce chormosomal damage may be taken as support for the premise that a BSDR results in transitions and, hence, cancer (NRC, 1988; Hall and Freyer, 1991).

A similar situation holds for links between the second level of causal evidence and the stage of neoplastic development, with several important differences. First, the mechanistic understanding of development is not as well developed as for conversion (see the discussions in Butterworth et al., 1989), leading to a weakened warrant via theory-based inference. While there is some limited agreement that development is related to changes in the kinetics of cellular mitosis, differentiation and replacement (see Chapter 2 and Cohen and Ellwein, 1990), there is much less agreement as to the particular biological structures (DNA membranes, etc.) requiring damage to yield the changes in kinetics. At least promotion (and to a lesser degree progression) has been associated with the onset of hyperplasia, with removal of hormonal control on cellular growth and differentiation, with changes in intercellular communication, and with extensive disruption of the histological architecture of cellular communities, but the steps leading to these changes have not been determined. As a result, the second level of evidence for neoplastic development will be limited to cases in which it is premised that development results from hyperplasia and lowering of intercellular communication. Hyperplasia often is indicated by observations on mitotic rates (Swenberg and Short, 1987) and/or labelling indices (SS), while interference with intercellular communication is indicated by changes in growth factor receptors and/or gap junction integrity (Trosko and Chang, 1988; Hartman and Rosen,

1983; Bertram, 1990). Changes in hormonal control may be indicated by observation of changes in hormone concentration, hormone receptor sites (structure or density on cell surface) or hormone structure, although these are not fully understood.

A feature shared by claims of neoplastic conversion and development is the need for premises embodying explicit mechanistic theories of carcinogenicity (unless the empirical correlation or existential insight warrant is invoked). The analyst must warrant use of a particular theory in linking the first level of evidence to claims of carcinogenicity, and the second level of evidence to claims of neoplastic conversion and/or development (see Section 5.2.3.). Considerations arising in the testing of theories were discussed in the appendix and are not repeated here. It is noted simply that the warrant for a particular theory can be (1) conceptual (the theory explains existing phenomena satisfactorily as judged by a suitably qualified expert), (2) empirical (the theory predicts the results of experiments) or (3) instrumental (the theory has allowed actions such as mitigation leading to desired outcomes such as lowered cancer incidence). When reflecting on these three forms of warrant for a theory, the distinction between verification and falsification made in the appendix is important. It also is necessary, when supporting a Theory-Based Warrant in Working Table 3, to examine the observational evidence under each of the existing theories and to determine the resulting coherence in the various theory-based inferences (see Section 5.5.).

5.2.5. *Warranting by Existential Insight*

At times, a warrant given by expert judgment will be available for a specific claim. In this case, there is no explicit line of reasoning leading from an observation to a specific claim. If there is a specific line of reasoning, this is included in one of the other relevance strategies or warrants of Working Table 3. This raises the issue as to whether expert judgments can be given

rational support. If rational support for a given instance of expert judgment cannot be offered, relevance strategies relying on existential insight will have weakened epistemic status.

The following premises constitute the rational basis of existential insight:

(1) It must be premised that the individual supplying the existential warrant has experienced the base of observations employed in the relevance strategy. Under existentialism, it is the physical setting of the observations which provides the necessary insight. Without existing in the physical setting, the individual expert cannot justify the necessary claim to experience from which insight (however unstructured it might be) arises.

(2) It must be premised that the individual possesses the necessary theoretical understanding to interpret experience. This is not to say that such theories are used explicitly to interpret the experience (in which case theory-based inference is the relevance strategy). Rather the theoretical understanding aids the expert to be "receptive" to the implications of the experience (in the language of existentialism, to "interpenetrate" with the observed phenomenon).

(3) It must be premised that the expert judgement has been elicited properly. Support for this premise has two components. First, the expert must be capable of retrieving the necessary insights from his or her psyche (meaning here whatever acts as the psychological or physiological site of the insight within the expert). The second condition is that the process of eliciting the judgment must not bias that judgment through specific wording or through the social/political/economic setting within which elicitations take place.

Finally, a brief note is in order concerning the strength of a warrant for existential insight. This strength rises as each of the three premises above receives support for a given expert. The first two premises (A and B) will be difficult to justify through reference to the psychological properties of the expert. The third premise (C) might receive some support by ensuring that the elicitation process satisfies established conditions of unbiased elicitation. Still, in most cases, the best support will be provided by the reliability of the expert in past instances (i.e. a form of correlation between past judgments and subsequent findings of carcinogenicity when more detailed information becomes available). As in all correlations, it must be established that a given instance of judgment is similar to the past instances on which the claim to reliability is established. If past instances of existential insight were based on a different

quality of observational evidence (see Chapter 3), or on a different theoretical grasp by the given expert (as when experts move outside their field of expertise), or by a different setting within which elicitations were obtained, then past reliability may not be an indication of present reliability by the expert. The key premise is that the present instance is characterized by the same conditions as underlie the set of past instances on which reliability is judged.

5.3. Warranting Inter-Context Premises

As described in the introduction to Chapter 5, the second role played by an observation lies in warranting any premises necessary to the use of relevance strategies in inter-context claims. These premises constitute the set of background assumptions necessary for assigning relevance to a particular observation with respect to a particular carcinogenicity claim extrapolated across contexts. The present section reviews briefly the role of specific observations in warranting specific background assumptions or premises. For a more detailed listing of the links between data categories and inter-context premises, the reader should consult Table 9 and the separate discussions in Section 5.4.

The specific premises to be warranted are:

(1) Exposure to a substance results in a BSDR both within the first (i) and second (j) contexts for which extrapolation is being attempted.

(2) Host factors and/or concurrent exposures (i.e. antecedent conditions) do not differ so significantly between the two contexts so as to suggest that a substance might be carcinogenic in one context but not the other.

(3) Cancers appearing at the first level of exposure are indicative of carcinogenicity at the second.

(4) Variability in exposure, pharmacodynamics and/or host characteristics within the first context (population) does not differ significantly from variability within the second context.

TABLE 9. ROLE OF DATA CATEGORIES IN INTER-CONTEXT PREMISES

Data Category[1]	Role of Data Categories in Inter-Context Premises
Tumor Response[2]	Can provide semi-empirical support for the premise that the relationship between BSDR and effect has a particular form, conditional upon establishing the premise of "Exposure to BSDR Conversion" by other data categories. If the above relationship has been established down to the BSDR of interest, the warrant is direct empirical. Also provides empirical correlation and existential insight warrants for the "BSDR to Effect Conversion". Does not provide support for the other premises.
Biophysical Effects	The data items have no bearing on the premises "Exposure to BSDR Conversion", "Host Factors", "Environmental Conditions" and "Intra(Inter) Subject Variability". They provide warrants for the "BSDR to Effect Conversion" premise by supporting the premise that a given BSDR of interest in extrapolation will be capable of inducing the effect. As discussed in "Broad Theoretical Implications", support for this premise depends upon whether the data item is (i) an indicator of cancer, (ii) an indicator of transitions between states of cancer of (iii) an indicator of conditions (mechanisms) necessary for transitions. Since none of these 3 indicators is an observation of cancer directly, this data category can provide only Theory-Based Inference, Existential Insight and/or Empirical Correlation warrants for the "BSDR to Effect Conversion" premise.
Host Characteristics	The data do not pertain to the premises on "Exposure to BSDR Conversion", "Environmental Conditions" and "Intra(Inter) Subject Variability". They may provide Direct Empirical and/or Theory-Based Inference support for the premise that "BSDR to Effect Conversion" is the same in two contexts by supporting the contention that (i) the target/mechanism is (or is not) present in both contexts, (ii) that necessary background transition rates for partial carcinogens do not differ significantly between the two contexts, and (iii) that repair of sublesions does not differ significantly between the two contexts.
Pharmaco-dynamics	The primary role of these data is in warranting the premise that there is (or is not) a significant difference in the relationship between exposure and BSDR for the two contexts, such that a BSDR might be produced in one context but not the other. This support will be Direct Empirical if the BSDR is measured directly in both species, Semi-Empirical if the BSDR is measured in both species but at higher (or lower) exposures than desired, or Theory-Based Inference if BSDR is not measured directly but other data items related to the production of a BSDR have been observed. If these "other" observations are correlated with production of a BSDR, Empirical Correlation warrants may be available.
Concurrent Environmental Conditions	The data provide support for the premise that there are (or are not) environmental conditions differing between 2 contexts that would call into question the causal role of a factor in the 2 contexts. The data also support the contention that Intra- and/or Intersubject Variability in exposure conditions does not differ significantly between the 2 contexts.

(continued on following page)

Table 9. (continued)

Data Category[1]	Role of Data Categories in Inter-Context Premises
Structure Activity Relationships	The data provide warrants (Theory-Based Inference, Empirical Correlation and/or Existential Insight) for the premise that 2 contexts do or do not differ significantly with respect to the ability of a BSDR to yield the effect. They provide this warrant by demonstrating that the substance is of a form capable of acting by a common mechanism in the 2 contexts, conditional upon the premise that a BSDR is produced in both contexts. The data also provide a warrant for the "Exposure to BSDR Conversion" premise by demonstrating that the substance is of a form capable of yielding an interaction (BSDR) in the two contexts, conditional upon the premise that a biologically significant burden is produced in both contexts.

1 See Working Table 2.

2 Quantitative measures are shown under "data items" for "tumor response" only. Quantitative measures commonly used are not given in other data categories, only the endpoints of interest.

These premises, and their warranting by specific data categories/items, are discussed in the separate sections of Section 5.4.

5.4. Linking the Observation Categories/Items to Premises Required by Inter-Context Relevance Strategies

The observational information within a context provides the most direct, and normally the most useful, source of evidence for warranting claims of carcinogenicity (warranting intra-context claims as discussed in Section 5.2.). For target contexts, however, one must resort to extrapolation of evidence from the observational contexts. Additionally, extrapolation to an observational context, from all other observational contexts, completes the utilization of all information available. The evidence from other observational contexts may serve to either strengthen or weaken the warrants for claims of carcinogenicity from the prior intra-context assessment. This step of completing the assessment for each observational context should precede extrapolation from observational contexts to target contexts. Working Table 4 is provided for data items of value in extrapolation across contexts. It is identical in format to Working Table 2, which is adequate for use in place of Working Table 4 if no additional data items specific to extrapolation need to be added.

Table 9 summarizes briefly the roles of various data categories in warranting inter-context premises. The relevance strategies available for extrapolation premises are displayed in Table 10 for reference. The link between specific observational evidence and inter-context premises is depicted in Working Table 5. The cells of this working table are discussed in separate sub-sections which follow. Assignments for the cells proceed in a manner identical to that employed in Working Table 3.

WORKING TABLE 4. DATA CHARACTERISTICS FOR OBSERVATIONAL CONTEXTS
FROM CONTEXT NO. ___ TO CONTEXT NO. ___

Data Category/Item	Completeness (Hi/Me/Lo/No)	Utility[1] (Hi/Me/Lo/No)	Observed Effect[2] (Hi/Me/Lo/No)			Causality[3] (AA/CC/EC/OC)
			Exposure Effect	Context-Specific Measurement	Organism-Specific Measurement	
Tumor Response TR1 TR2 .						
Biophysical Effects BE1 BE2 .						
Pharmacodynamics PD1 PD2 .						
Host Characteristics HC1 HC2 .						
Related Substances Assessments RSA1 RSA2 .						

(continued on following page)

Working Table 4 (continued)

Data Category/Item	Completeness (Hi/Me/Lo/No)	Utility[1] (Hi/Me/Lo/No)	Observed Effect[2] (Hi/Me/Lo/No)			Causality[3] (AA/CC/EC/OC)
			Exposure Effect	Context-Specific Measurement	Organism-Specific Measurement	
Structure Activity Relationships SAR1 SAR2 . . .						

1 May be subdivided into additional categories when useful, e.g. "validity", "reliability", and "accuracy" may be judged separately for data items from laboratory sources.

2 Refers to effect on data item of exposure to agent. Footnotes with explanatory comments may be needed. When statistical measures are available, they may be more informative than a simple indication of Hi/Me/Lo/No, e.g., estimates or tests of statistical significance.

3 See Section 4.2.4 for explanation of the following choices available.

AA: Accidental Association
CC: Common Cause
EC: Empirical Causality
OC: Operational Causality

TABLE 10. RELEVANCE STRATEGIES AVAILABLE FOR EXTRAPOLATION PREMISES

Data Category	Exposure to BSDR Conversion	BSDR to Effect Conversion	Host Factors	Environmental Conditions	Intra(Inter)-Subject Variability
Tumor Response	N.A.	D.E. S.E.E. E.C. E.I.	N.A.	N.A.	N.A.
Biophysical Effects	N.A.	T.B.I. E.C. E.I.	N.A.	N.A.	N.A.
Host Characteristics	N.A.	N.A.	D.E. T.B.I. E.C. E.I.	N.A.	N.A.
Pharmaco-dynamics	D.E. S.E.E. T.B.I. E.C. E.I.	N.A.	N.A.	N.A.	D.E. T.B.I. E.I.
Concurrent Environment Conditions	N.A.	N.A.	N.A.	D.E. T.B.I. S.E.E. E.C. E.I.	D.E. S.E.E. T.B.I. E.C. E.I.
Structure Activity Relationships	T.B.I. E.C. E.I.	T.B.I. E.C. E.I.	N.A.	N.A.	N.A.

[1] N.A.: None Available
D.E.: Direct Empirical
S.E.E.: Semi-empirical Extrapolations
E.C.: Empirical Correlations
T.B.I.: Theory-based Inference
E.I.: Existential Insight

WORKING TABLE 5. SUPPORT FOR INTER-CONTEXT EXTRAPOLATION PREMISES
CONTEXT NO. __ TO CONTEXT NO. __

Reference Strategy	Exposure to BSDR Conversion	BSDR to Effect Conversion	Host Factors	Environmental Conditions	Intra(Inter)-Subject Variability
Direct Empirical (D.E.)					
Semi-Empirical Extrapolation (S.E.E.)					
Empirical Correlation (E.C.)					
Theory-based Inference (T.B.I.)					
Existential Insight (E.I.)					

Overall Assessment					

Working Table 6 is provided for the assessment for inter-context support for claims of carcinogenicity. That table is identical in format to Working Table 3 for assessment of intra-context assessment. Working Table 4 (on data characterisitics for inter-context extrapolation) is utilized in the same way for completion of Working Table 6 as Working Table 2 (on data characteristics for observational contexts) is utilized for completion of Working Table 3. The judgments in Working Table 6, however, should be based on the support for inter-context extrapolation premises depicted in Working Table 5, as well as the data characteristics depicted in Working Table 4. The intra-context assessment (applicable to observational contexts but not target contexts) and the inter-context assessment (applicable to both observational and target contexts) are depicted and then summarized in Working Table 7 for overall assessment from all sources.

5.4.1. *Conversion from Exposure Conditions to BSDR*

As stated in the broad theory of environmental carcinogenesis, it is assumed that a substance in the environment must produce a dose-rate of the biologically significant form of the substance within a target organ, tissue, cellular subpopulation, etc. It must be demonstrated, therefore, that there are no important differences between the study population constituting the observational data and the population for which claims to carcinogenicity will be made, at least with respect to the relationship between exposure and BSDR. By "important" is meant differences which would strengthen (or weaken) a claim that cancers in the study population might be due to pharmacodynamic factors present in that population but not in the population for which claims of carcinogenicity will be made (presumably, the target context). The inverse of this issue may also apply, in which case reflection on pharmacodynamic factors might suggest that an observation of "no cancer" in the study population might be due to factors present (or absent) in that population but not in the population for which claims to carcinogenicity will

WORKING TABLE 6. INTER-CONTEXT SUPPORT FOR CLAIMS OF CARCINOGENICITY[1]
FROM CONTEXT NO. ___ TO CONTEXT NO. ___

Reference Strategy	I.O.[3]	Claims of Carcinogenicity[2]: Increases Incidence of Cancer	Classification(s): Complete	Partial	Mixer	Helper	Stage(s): Neo. Conv.	Neo. Devel.	Mechanism(s): Geno-toxic	Non-genotoxic
Direct Empirical (D.E.)										
Semi-Empirical Extrapolation (S.E.E.)										
Empirical Correlation (E.C.)										
Theory-based Inference (T.B.I.)										
Existential Insight (E.I.)										
Column Summary										
Overall Summary										

1 Top half of each entry is completed using dose-response data observed in the context. Bottom half of each entry includes "floater-data" as well (see text).

2 Choices for cell entries are Hi/Me/Lo/No. See text of Section 5.1. This assignment is made independently of the assignment of intellectual obligation.

3 "Intellectual Obligation" (Hi/Me/Lo/No). See text of Section 5.1. This assignment is made independently of the assignment in the cells of the table (footnote 2).

WORKING TABLE 7. SUMMARY OF OVERALL ASSESSMENTS
FOR CLAIMS OF CARCINOGENICITY BY CONTEXT
CONTEXT NO. __

Context Number	Claims of Carcinogenicity[1]								
	Increases Incidence of Cancer	Classification(s)				Stage(s)		Mechanism(s)	
		Complete	Partial	Mixer	Helper	Neo. Conv.	Neo. Devel.	Genotoxic	Non-genotoxic
Intra-Context[2] No. __									
Inter-Context[3] No. __									
Inter-Context[3] No. __									
Inter-Context[3] No. __									
Inter-Context[3] No. __									
Column Summary									
Overall Summary									

1 Choices for cell entries are Hi/Me/Lo/No.
2 Context number should match context number in table heading.
3 Number of the context from which results were extrapolated.

be made. In that case, observation of no carcinogenicity in the study population does not warrant a claim that the substance is a non-carcinogen in the target context. These two possibilities of judgment are not distinguished further in this report.

For the case of radon, differences in the relationship between exposure and BSDR within the mining population (i) and home population (j) were explored at length in the second report on radon (NRC, 1991). These differences resulted from population-specific differences in degree of attachment of radon progeny to aerosols, aerosol size distribution, equilibrium fraction, breathing characteristics, lung sizes, and mucus flow rates in the lung. The resulting analysis warranted the claim that the BSDR generally would be lower in the present following exposure to airborne radon. This conclusion was warranted regardless of the assumed location of target cells. The primary uncertainty, then, was whether this lower BSDR was capable of inducing cancer (see Section 5.4.3.).

Radon is an extreme instance in the sense that pharmacodynamic data are plentiful. It must be noted that complete understanding of the pharmacodynamics for a substance often is lacking. This may arise if (1) pharmacodynamic properties are not understood (2) the target for action is unknown or (3) the biologically significant form of the substance is unknown. Reference to Figure 1 is useful here in determining how pharmacodynamic data may still be useful even when one of the above 3 cases of weakened understanding applies. From this figure, it may be seen that the chain of reasoning in pharmacodynamics proceeds from exposure to intake to uptake to organ burden to biologically significant organ burden to biologically significant dose-rate (BSDR).

This introduces the following premises into the analysis in supporting the judgments on the relationship between exposure and BSDR (the premises are given first in sketch, but are discussed in more detail at the end of this section):

Premise 1: There are no differences between the study population (hereafter, P_S) and the population of interest in making claims to carcinogenicity (hereafter, P_I), with respect to the conversion from exposure to intake, which would detract from the claim that cancers in P_S imply cancers in P_I at equal levels of exposure. The

factors influencing this conversion are (a) inhalability of the physical form of the substance in the environment, as when filters are present or particle sizes in air are incapable of entering the nose and/or mouth; (b) breathing rates, ingestion rates or degree of skin contact (for cases of inhalation, ingestion and dermal exposures, respectively); (c) location or presence of P_S and P_I within the field of exposure. Differences in these factors (discussed at length in NRC, 1991) may arise either from inherent biological differences between P_S and P_I (as in interspecies extrapolation) or from differences in exposure to other substances that modify the relationship between exposure and intake. The same two sources of difference apply to the other 4 premises that follow. The analyst reflects on these issues and judges whether intake in P_S will be higher or lower than the intake in P_I.

Premise 2: There are no differences between P_S and P_I, with respect to the conversion from intake to uptake, which would detract from the claim that cancers in P_S imply cancers in P_I at equal levels of intake. The factors influencing this conversion are (a) deposition fraction in the lung (for inhalation), (b) absorption from the lung to the body fluids (blood, water, etc), (c) absorption from the body fluids to the target organ, (d) first-pass excretion, (e) dermal absorption (for dermal exposures) and (f) absorption from the G.I. tract to the body fluids (for ingestion). If the target organ is the G.I. tract, then intake and uptake are identical for the ingestion route of exposure. The same applies to the lung for inhalation exposures. The analyst reflects on these issues and judges whether organ uptake in P_S will be higher or lower than the uptake in P_I (see NRC, 1991).

Premise 3: There are no differences between P_S and P_I, with respect to the conversion from uptake to organ burden, which would detract from the claim that cancers in P_S imply cancers in P_I at equal levels of uptake. The factor influencing this conversion is the retention of the substance in the organ. The analyst reflects on this issue and judges whether organ burden in P_S will be higher or lower than the organ burden in P_I (see NRC, 1991).

Premise 4: There are no differences between P_S and P_I, with respect to the conversion from organ burden to biologically significant organ burden, which would detract from the claim that cancers in P_S imply cancers in P_I at equal levels of organ burden. The factor influencing this conversion is the fraction of the original substance transformed into the biologically active form. The analyst reflects on this issue and judges whether biologically significant organ burden in P_S will be higher or lower than the biologically significant organ burden in P_I (see NRC, 1991).

Premise 5: There are no differences between P_S and P_I, with respect to the conversion from biologically significant organ burden to BSDR within the target organ. Factors influencing this are (a) the presence of target sites within the organ, (b) spatial location of the substance molecules with respect to the target sites and (c) ability of the substance molecules to interact with the target sites. It should be noted that this fifth premise rarely can be supported due to a lack of information on the target site (an exception being radon (NRC, 1991)). Still, the analyst reflects

on this issue and judges whether BSDR in P_S will be higher or lower than the BSDR in P_I.

Reflecting on each of the five premises, the analyst assigns an index of evidential support to each premise (scale of low to high). If all five premises are well supported by evidence, with the evidence being given by the factors discussed under each premise above, the judgment under the "Exposure to BSDR Conversion" premise of Working Table 5 is deemed strongly supported and is assigned an overall index of "high". As support for any of the premises is weakened, the index on the judgment falls towards "low".

Support for a given premise may be in the form of any of the 5 relevance strategies of Section 5.1. For example, consider Premise 3 (concerning the relationship of uptake to organ burden). There may be observational data available on the ratio of organ burden to organ uptake in both P_I and P_S. In that case, there is a direct empirical warrant for claims about this ratio, conditional on the observations having been obtained at the uptakes of interest in the analysis. If the uptakes are larger than those of interest in the analysis (such as when retention functions are obtained from acute and large uptakes), the observed ratio must be extrapolated to lower levels of uptake. In that case, the warrant for premise 3 is semi-empirical. If information on the ratio is not available, but it is deemed that the ratio in P_I generally correlates well with the ratio in P_S (perhaps in examinations of other substances for which observational evidence is available), then a warrant of empirical correlation is made. If the analyst simply judges that the ratio should be similar in P_I and P_S, without giving explicit observational warrants, the result is a warrant of existential insight. Finally, if direct observations of the ratio are unavailable, but observations on the retention function are available, the analyst may judge the ratio to be the same (or different) in P_S and P_I based on the similarity (or difference) in retention. This warrant is theory-based inference since it does not involve a direct measurement of organ burden in either P_I or P_S and requires a further premise that the identified important factors (here, retention) are the only ones significantly affected the ratio (organ burden to uptake) in P_I and P_S.

The preceding paragraph focused on the five premises necessary to judge the strength of the warrant for the "Exposure to BSDR Conversion" Premise in Working Table 5. Having reflected on these five premises, the analyst determines which of the five relevance strategies may be applied to the "Exposure to BSDR Conversion" premises and assigns an indication of weight-of-evidence (Hi to Lo). The forms of warrant for any of the inter-context premises appear explicitly in Working Table 5. It will be noted, however, that the premises necessary to support a given inter-context premise (such as premises 1 through 5 above) are not depicted explicitly in this working table. As a result, the form of the warrant for premises 1 through 5 above is used in the judgment of the strength of the inter-context premise (Working Table 5), but does not appear explicitly in Working Table 5.

In addition, it may be noted that it is not necessary to warrant each of the five premises from this section separately in order to warrant the inter-context premise (concerning the relationship between exposure and BSDR in the two contexts). Any of the five premises may be combined if direct observations on the relationship between any step in the analysis of the judgment (exposure-intake, intake-uptake, etc.) and any "earlier" step is available. For example, measurements of the relationship between exposure and organ burden may preclude the need for separately warranting premises 1 through 3. The observational data warrant the three premises (1 through 3) simultaneously by showing that the ratio of organ burden to exposure is similar (or different) in P_S and P_I. The strongest conceptual strength of the warrant for the three premises remains, of course, one in which all three premises are warranted separately, giving support to the claim that (1) the ratio is similar (or different) in P_S and P_I and (2) the similarity (or difference) is understood in terms of the analytic steps leading to an estimate of organ burden. Still, even if the reason for the similarity (or dissimilarity) is unexplainable, there may exist direct measurements warranting premises 1 through 3 empirically.

Finally, a comment is in order concerning the role played by the above premises in the judgment that the inter-context premise is warranted. Unless a threshold BSDR necessary for carcinogenicity or

a non-monotonic relationship between BSDR and effect is premised, any value of BSDR observed to produce cancer in P_S provides a warrant for the claim that the substance is a carcinogen in P_I (conditional upon the "applicability" of the antecedent conditions in P_I). This applies regardless of the numerical value of the BSDR in P_I. The judgment that consideration of factors concerning conversion from exposure to BSDR does not alter the claim of carcinogenicity in the analysis, therefore, hinges on the ability to demonstrate only that population P_I will receive at least a non-zero BSDR. This will be true so long as it can be shown that:

(1) Intake occurs in P_I, only requiring premises that (1) the population is exposed to the substance in the environment, (2) does not have a filter of perfect efficiency placed over the lung, G.I. tract or skin (depending upon the route of exposure), and (3) the lung, G.I. tract or skin does not block movement of the substance into the body completely (as when the rat nasal passages close down during exposure to high concentrations of formaldehyde (Graham et al., 1988)).

(2) The uptake fraction is not zero in the target organ, only requiring premises that (1) there is not complete first-pass excretion of the substance, (2) the substance is not completely exhaled, regurgitated, or washed from the skin (depending upon the exposure route) and (3) uptake into the organ does not require saturation of some mechanism of absorption.

(3) The substance is not removed immediately from the target organ upon entry, requiring only the premise that the retention function is non-zero.

(4) The substance is metabolized to some degree into the active form, requiring premises that (1) this metabolic process be present in P_I and (2) that the process is not characterized by saturation. If metabolism is from the inactive to the active form), satisfying premise (2) requires that the activation process be operative at the organ burden, expected in P_I and not be set in motion only by higher burdens. If metabolism is from the active to inactive form, satisfying premise (2) requires that the inactivation process be less than 100% efficient at the organ burden expected in P_I.

(5) A target (organ, tissue, etc.) is present in both P_S and P_I. This will be warranted most strongly when the target has been identified, but in many cases the target is not known and must be assumed to be present in both populations based on anatomical similarities.

5.4.2. *Host Factor and/or Concurrent Conditions Do Not Differ Significantly*

From the judgment discussed in 5.4.1., the analyst determines that both P_S and P_I are receiving some non-zero value of BSDR (with an associated index of evidential support). This supports the judgment that targets of individuals in P_I are being dosed, to some degree, by the active form of the substance in the target organ (tissue, cell, etc.). It might then be assumed that claims of an effect in P_S (presumably due to the presence of a BSDR at the target site) is warrant for a claim that an effect will occur in P_I (where a BSDR also has been judged present). This assumption, however, is based on the premise that the causal association between BSDR and effect in P_S is not due to conditions other than the BSDR that might be present in P_S but not P_I. These conditions are antecedent to exposure to the substance under analysis in the sense that they establish the biological and exposure conditions within which the substance under analysis exerts an effect. For example, in the case of ETS, it was argued that the Chinese studies were not applicable to the U.S. population due to the presence of large background levels of cooking and heating smoke in homes. A similar argument was made for radon, in which it is was claimed that the large dust burden in mines allowed expression of the carcinogenic properties of radon (see the discussion in NRC, 1988). It was claimed further that this antecedent condition related to concurrent exposures (presence of large dust burden) was present in mines but not in homes, calling into question the premise that cancer in the mining population (P_S) was due to the inherent action of radon in the lung rather than to the antecedent condition (dust).

The discussion in this section is directed towards warranting a judgment that an effect produced by the substance of interest in P_S did not require the simultaneous presence of antecedent conditions (either concurrent exposures or host factors) in P_S unlikely to be present in P_I. This judgment clearly is related to the judgment that a substance is (or is not) a "mixer" or "helper". The inverse of this judgment is that a lack of effect produced by the substance of interest in P_S was not the result of lacking the necessary antecedent conditions in P_S, with those antecedent conditions being present in P_I. The

judgment under discussion here presupposes that the substance played a causal role in the effects noted in P_S, and focuses on the judgment that this causal role was (or was not) conditional upon antecedent conditions present in one population (P_S or P_I) but not the other.

The role of antecedent conditions appears in the theory of carcinogenicity employed in an analysis. It must be premised that there is nothing present in the exposure conditions and/or in bodies of individuals in P_S, but missing in P_I (or vice-versa), that would significantly alter the relationship between BSDR and carcinogenicity. These antecedent conditions might relate to the following factors:

(1) Initial state vector (Crawford-Brown and Hofmann, 1990). Prior to exposure to the substance of interest, P_S and P_I might differ in the degree to which the target already has been initiated and/or promoted. This could arise from differences in previous exposures or from host factors. For example, exposures to high dust levels in mines have been suggested to establish promotional action in lungs of miners (P_S) that would not be present in homes (P_I). If it is assumed that (1) radon does not induce both initiation and promotion in P_I and (2) that there are no sources of promotion in P_I other than exposure to dust, then carcinogenicity from radon in P_S might be claimed to be due entirely to the antecedent conditions of promotion brought on by high dust exposures. This weakens the claim that radon is a carcinogen in the home environment where dust levels may be in sufficient to induce promotional action.

The initial state vector also is important when it is premised that a target (organ, tissue, etc.) must possess a minimal number of cells in a given state (initiation, promotion and/or progression) to yield a fatal tumor. If the number of cells in that state already is sufficient, prior to exposure to a given substance, that substance will have no effect on the number of cells in the considered state. If the prior number of cells is below, but close to the threshold number, exposure to the substance might increase the incidence of cancer by inducing transitions and raising the number of cells (in the given state) above the threshold. If the prior number of cells in the considered state is well below the threshold, the substance may yield an insufficient number of transitions to exceed the required threshold. All of the above cases, of course, are significant if, and only if, a threshold number of cells in a given state is required.

(2) Repair rates and/or efficiency. Particularly in the case of initiation (Hall and Freyer, 1991), it has been established that initial damage produced by a carcinogen can at times be repaired. Two populations (P_S and P_I) may differ in the degree of repair for this initial damage. For example, individuals with xeroderma pigmentosum lack the necessary enzymes for operation of the DNA repair system, resulting in greatly increased sensitivity to UV induced skin cancer.

Reflection on repair rates requires consideration of several other factors. If repair is 100% efficient in P_S but not in P_I, then a finding of no carcinogenicity in P_S does not fully warrant a claim of no carcinogenicity in P_I, conditional upon the premise that the

substance of interest plays a role in carcinogenicity mediated by a process subject to repair. If repair is 100% efficient in P_I but not in P_S, then a finding of carcinogenicity in P_S does not fully warrant a claim of carcinogenicity in P_I, again conditional upon the premise that the substance of interest plays a role in carcinogenicity mediated by a process subject to repair. If repair is less than 100% efficient in both populations, it must be established that repair does not lower the number of unrepaired damage sites below some threshold value. Observations relevant to the issue of repair are measurements of repair rates.

(3) Background transition rates. These rates refer to the rates of transition between normal, converted and developed cells without the presence of the substance of interest. This consideration might be important due either to biological differences in P_S and P_I (i.e. host factors) or differences in concurrent exposures. The significance of this factor arises when it is possible to saturate processes leading to a given transition. For example, consider the case of a substance that acts through stimulation of mitosis, yielding a transition from initiated to promoted states. If the mitotic rate already is saturated in a given population (P_S or P_I), exposure to the substance of interest might prove incapable of increasing the incidence of tumors. In another population with an unsaturated mitotic rate, exposure to the substance might increase the mitotic rate and, hence, the incidence of cancer. It should be noted here that the issue being raised is not one of the numerical value of the background transition rate but whether the transition rate can be further modified by action of the substance of interest.

Another case where consideration of background transition rates are important is when an etiologic process must exceed a threshold to yield the transition. For example, it was suggested in the case of formaldehyde that promotion brought on by induction of hyperplasia requires a threshold level of hyperplasia (Swenberg et al., 1983, 1987). A given population (P_S or P_I) when exposed to a hyperplastic agent might possess a sufficiently low background degree of hyperplasia that the added hyperplastic action of the substance of interest does not increase hyperplasia above the threshold required for a transition to the promoted state. A second population might possess a background degree of hyperplasia close to the threshold, so that the added hyperplastic action of the substance of interest does increase hyperplasia above the required threshold.

(4) Presence and/or absence of target in host.

As with all of the considerations in this section, the differences in background rates of transition may be due to differences in the concurrent exposures or inherent biological properties of a population.

One further premise must be introduced in employing measured or estimated background rates of transition. In its simplest form, the multistage theory of carcinogenesis employs stages, and transitions between those stages, which are common to all carcinogens. For example, it might be assumed that all instances of neoplastic conversion share common events leading to neoplastic

development. This theoretical approach to carcinogenesis presupposes a general structure to carcinogenesis that is invariant between specific carcinogens. To the degree that carcinogens act in different ways to produce different forms of conversion and development, the premise of commonality will be invalid. Measurements of background rates of transition may require selection of only those forms of transition germane to the specific carcinogen under study. It must be premised, therefore, that the background transition rates selected in an analysis represent the rates applicable to the specific forms of conversion and development required by the specific carcinogen within the contexts under consideration for extrapolation.

5.4.3. *Extrapolation Across Doses and/or Dose-Rates*

One of the most contentious areas of hazard identification concerns the judgment that observation of tumor response carcinogenicity at high values of the dose (biologically significant dose) and/or dose-rate (BSDR) provides a warrant for claims to carcinogenicity at values of interest in the regulatory setting. If the task of hazard identification is to warrant the claim simply that a substance is a carcinogen at some level of dose (BSDR, etc.), then observations of carcinogenicity under high levels of dose (BSDR, etc.) serve as an appropriate warrant for that claim. If, however, the task is to determine whether the substance is a carcinogen within a context of interest to the regulator (with these conditions typically being characterized by relatively low values of dose, BSDR, etc.), then premises concerning the effect of lowering doses (BSDR, etc.) below those contained in the "high exposure" context must be introduced into the analysis and warranted. It is towards elucidation of these latter premises and their warrants that the present section is directed. There may, of course, be observational evidence available for the target context, but use of the evidence constituted the intra-context warrants discussed previously.

The strongest warrant for the premise that causal relationships between BSDR and tumor prevalence are constant (but not numerical identical) is a direct empirical observation that the BSDR is P_S and P_I produces a response in both populations. This requires tumor response data at doses (BSDRs) encompassing both P_S and P_I contexts. Such data might arise, for example, from injection (uptake) studies in which the relationship between exposure and uptake is not at issue. It must, of course, be premised that route of administration (injection rather than environmental exposures) does not alter the relationship between BSDR and response in a manner which differs between large and small uptakes.

If the available tumor response data do not encompass dose-rates encountered in P_I, semi-empirical warrants supplant the direct-empirical warrants for the premise under consideration in this section. In this case, the claim of carcinogenicity at high dose or dose-rate is extrapolated to low dose or dose-rate based on direct observation of a pattern in the relationship between BSDR and response in P_I. There are two routes to semi-empirical warrants. The first is a warrant of observed patterns in a plot of dose versus tumor response (dose-rate being held constant) and/or a plot of dose-rate versus tumor response (dose being held constant) for the population (context) P_S. The analyst judges that the pattern has been directly observed in the data on P_S. This pattern then is followed visually to the dose and/or dose-rate of interest in P_I, supporting (or weakening) the contention that the substance is a carcinogen in P_S at low dose and/or dose-rate. The key premise here is that the pattern is observationally evident in the data rather than being imposed on the data by a fitting equation. This requires, of course, tumor response data of sufficiently quality to bring any underlying patterns into view for the analyst, both for the case of dose and dose-rate. Such data were available for the case of radon exposures, since the mining populations could be divided into groupings characterized by different doses and dose-rates (NRC, 1988). It should be noted, however, that patterns in the data were not clear, requiring the addition of a second form of warrant discussed in the following paragraph.

The second warrant for semi-empirical extrapolation is theory-based semi-empirical extrapolation. In this case, the same observational data discussed above (on dose and dose-rate versus tumor response) are employed, but the requirement of direct observation of patterns in the data is dropped. Instead, a fitting equation deduced from etiologic theory is employed in the extrapolation. The NRC committee chose to employ a linear dose-response equation based on belief in a one-hit model of radiation carcinogenicity (NRC, 1988). The extrapolation equation is not used in hazard identification to estimate the actual risk at low exposures, but simply to determine whether carcinogenicity is to be expected at lower doses and/or dose-rates. In addition, the premise that the extrapolation equation itself is valid must be warranted by warranting the premises appearing in the etiologic theory from which the equation is deduced (again, see the discussion in Section 5.2.). The strongest case of semi-empirical extrapolation will hold when both (1) the pattern is evident in the data and (2) the pattern is to be expected from the prior establishment of an etiologic theory and its associated extrapolation equation.

Both the empirical correlation and existential insight warrants for extrapolation were discussed previously (see Sections 5.2.3. and 5.2.5., respectively). That discussion will not be repeated here, other than to comment that the correlation is between a finding of carcinogenicity at high doses and/or dose-rates and carcinogenicity at low doses and/or dose-rates. This correlation might be expected to improve as the dose and dose-rate in P_S approach the values in P_I. The analyst may choose, therefore, to construct the measure of correlation on sets of substances for which the magnitude of difference in dose and/or dose-rate is similar to that of interest in the analysis.

Finally, the warrant for extrapolation may be in the form of theory-based inference. Only data items from data categories other than "tumor response" may be employed here, since the latter were employed in the direct empirical and semi-empirical warrants. In this case, the most explicit use of etiologic theories is made in developing the necessary premises. Again, both dose and dose-rate must

be considered. The remaining discussion in this section focuses on consideration of differences in dose and dose-rate between P_S and P_I.

With regards to both dose and dose-rate, four primary premises must be warranted. These are:

(1) There is not a dose and/or dose-rate below which the transitions produced by the substance of interest will not occur (or at least the dose and/or dose-rate in both P_S and P_I exceed the threshold). For example, it has been proposed that DNA repair mechanisms can deal effectively with initiating damage so long as the rate of damage is not sufficient to induce SOS repair. It also has been proposed that hyperplasia-induced promotion by formaldehyde requires a threshold dose-rate. It also has been proposed that a minimal level of organizational disruption in cellular communities is required to induce cancer. This disruption presumably is a function of both total dose and dose-rate.

(2) There are not competing processes which change their "order" of impact on carcinogenicity below some threshold level (or at least the doses and/or dose-rates in P_S and P_I are either both above or both below this threshold). For example, many substances both induce transitions to states of cancer and are cytotoxic (for an application to the case of radon, see Crawford-Brown and Hofmann, 1990). Substances which induce transitions will tend to be carcinogenic. Substances which are cytotoxic may kill cancerous cells, thereby lowering the incidence of cancer in a population. It might be the case that a substance induces newly cancerous cells at a rate faster than it kills previously cancerous cells, leading to a net observation of carcinogenicity at a given dose and/or dose-rate in P_S. At lower (or higher) doses and/or dose-rates, however, cytotoxicity to previously cancerous cells may exceed induced transitions, leading to a net observation on non-carcinogenicity (or even "life-saving") at these lower (or higher) doses and/or dose-rates.

(3) The latency period for cancer does not increase significantly at low doses and/or dose-rates. For example Raabe et al. (discussed in NRC, 1988) have proposed that the latency for bone sarcoma might exceed the expected lifetime at low doses and/or dose rates. If this is the case, the latency for cancer in P_S might be less than the expected lifetime (leading to an observation of carcinogenicity), but the latency for cancer at lower doses and/or dose-rates in P_I might exceed the expected lifetime (leading to an observation of non-carcinogenicity).

(4) There is not a mechanism for carcinogenesis present (or absent) at the level of dose and/or dose-rate in P_S but not present (or absent) at the level of dose and/or dose-rate in P_I. If this mechanism simply requires exceedance of a threshold for one of the transitions to cancer, then premise 4 and premise 1 are identical.

5.4.4. *Consideration of Intrasubject and Intersubject Variability*

Epidemiologic and experimental studies of the effect of an environmental substance typically employ estimates of the mean exposure, dose, etc. in defined groups (such as exposure groups in a cohort study). The result is an estimate of the relationship between mean level of exposure (or dose, etc.) and the response in a population.

Such approaches presuppose that response is predicted entirely by mean exposure rather than by other properties of the distribution of exposure within a group characterized by inhomogeneity of (1) exposure and (2) sensitivity to the action of a substance. In some cases, however, knowledge of the mean exposure in a group is not sufficient to determine the response of the group (for a discussion specific to radon, see Crawford-Brown and Hofmann, 1989) either qualitatively or quantitatively. In such cases, other properties of the distribution such as variance may be of equal importance in predicting response. This section summarizes the important consideration of variability within an exposed population.

The variability has two primary components: intrasubject and intersubject variability. Intrasubject variability refers to variations in exposure (intake, uptake, burden, dose, BSDR, etc.) both spatially and temporally with respect to an individual organism such as a human or experimental animal. This variation arises from the following factors:

(1) Exposure conditions may vary for the individual, as when radon concentrations in home air fluctuate during the day. If it is premised that BSDR must exceed a threshold value to produce cancer, and only if this is premised, the analyst must determine whether temporal variation in exposure conditions affects the judgment of carcinogenicity. For example, the study population (P_S) may have been exposed to a well controlled environment where the mean exposure (and, hence, BSDR) was below the required threshold for cancer. The population of interest (P_I) may be exposed to the same mean exposure but with greater temporal variation. As a result, cancer may not be present in P_S but might be expected in P_I if the variation in the latter exposure produced intervals of time during which the BSDR exceeded the threshold. Conversely, excessive variation of exposures in P_S might produce cancer while none is expected in P_I due to lesser variation. The key issue is whether the extent of variation differs between P_S and P_I to such a degree that thresholds of BSDR are exceeded in one population but not the other.

(2) The BSDR may vary spatially within the target organ of an individual due to inhomogeneity of uptake or retention (see the information in Crawford-Brown and Hofmann, 1989). Again, this spatial variation in BSDR may cause some fraction of the organ to receive a BSDR above the threshold required for cancer (requiring, of course, a premise that a threshold exists). If the inhomogeneity of BSDR within the target organ differs dramatically between individuals in P_S and those in P_I, cancer might be caused in one but not the other.

The second source of variability is intersubject variation. As in the case of intrasubject variation, intersubject variability is significant to the task of hazard identification only if thresholds for cancer are premised. The important sources of intersubject variability are as follows:

(3) Exposure conditions may vary between individuals in a group. This variability is complicated by variability in pharmacodynamic properties for those individuals. The result of these two factors is variability in BSDR between individuals in the group. Again, premising a threshold BSDR necessary for cancer, the determinant of response in the group is not mean exposure or mean BSDR but the fraction of individuals with a BSDR above the threshold. A finding of cancer (or no cancer) in P_S does not, therefore, warrant a claim of carcinogenicity in P_I if the variability within P_S and P_I differs to such a degree that the fraction of individuals exceeding the threshold is zero in one population but not the other.

(4) Individuals may vary with respect to sensitivity (Redmond, 1981). This variation is significant for hazard identification if (and only if) sensitivity is characterized by a threshold BSDR necessary for cancer. In that case, it must be premised that variability of sensitivity in P_S and P_I do not differ to such a degree that there are (1) individuals in one population with a threshold below the delivered BSDR but (2) not in the other population. This consideration gains significance when P_S is constituted by genetically similar experimental animals (presumably with similar thresholds) and P_I is constituted by a diverse human population (presumably with variation in thresholds, if thresholds exist). It has been argued, for example, that a finding of a threshold for cancer in a genetically homogeneous population does not warrant a claim of non-carcinogenicity in human populations characterized by a wide range of sensitivities.

5.5. Column/Row Summaries and the Issue of Coherence

The various working tables display a number of points at which the issue of coherence must be raised in order to provide the entries for specific cells. By coherence here, we mean the degree to which a set of observations, theories, inferences, etc., present a unified and supporting pattern of

warrant for a specific claim. Coherence is lost whenever there are important differences between claims of carcinogenicity drawn on the basis of different:

(1) Cases within a data item.

(2) Data items within a data category.

(3) Theories within "Theory-Based Inference".

(4) Relevance strategies within a context.

(5) Contexts.

These issues, or bases of difference, are discussed separately in the present section.

First, a distinction between two aspects of coherence must be drawn. The analyst first should determine whether the AVAILABLE evidential base displays coherence. This is referred to here as extant coherence since it is a judgment that the existing evidential base leads to a consistent (non-contradictory) judgment of carcinogenicity. Of equal importance, however, is a second form of coherence referred to here as complete coherence. Complete coherence arises when all data items potentially of use in analysis (in any manner of use) are (1) available and (2) present a consistent, mutually supportive, pattern of evidence. For example, a number of epidemiologic studies might be available (as in the direct empirical studies of radon in the home), leading to a claim of extant coherence if their results are similar (this was not the case for the direct empirical studies of radon but was the case for the semi-empirical mining data). Still, there may be greater support available if pharmacodynamic data displayed the existence of a BSDR and if biophysical effects were observed (as in the case of radon exposures). The existence of these latter data, if they yield similar inferences of carcinogenicity (as they do with radon), strengthen the coherence by increasing the claim to complete coherence.

Examining Working Tables 2 through 7, then, it may be noted where issues of coherence arise and must be factored into assignments for cells appearing in those tables. For Working Table 2, the

coherence of the cases under a given data item is an essential component of the judgment of "observed effect" and of "causality" (see the previous discussion in Section 4.2.). For Working Table 3, coherence arises in several instances. Within a relevance strategy, different data items may be pertinent to a given strategy due to the existence of several premises required for invocation of that strategy (see Section 5.2.). If the premises are supported consistently by the available data, extant coherence may be high. Complete coherence will be high only if all data potentially of use in supporting the premises are available and yield a consistent judgment concerning the premises. Coherence also enters in the use of etiologic theories within the "Theory-Based Inference" cells. Here, the analyst must determine if a similar claim of carcinogenicity is implied by each of the existing theories. Coherence in this instance rises from "Lo" to "Med" to "Hi" as the various theories move the judgment form one of "inconsistent inference between examined theories", to "consistent inference between examined theories" (extant coherence) and, finally, to "complete coherence between a large body of theories" (complete coherence). In this regard, it is of interest to note that etiologic theories of radon carcinogenicity do not uniformly lead to claims that semi-empirical extrapolation suggests carcinogenicity at lower exposures in the home.

Within a claim in Working Table 3, the analyst then examines the coherence between inferences drawn on the basis of different relevance strategies. If all of the available strategies yield a consistent inference, an instance of extant coherence applies. If, in addition, all relevance strategies were available (as is true for radon), complete coherence applies. When inconsistency between strategies exists, the analyst must examine the claim to "intellectual obligation" to determine the impact of this incoherence on the epistemic status of a claim. Incoherence then weakens the claim only if the incoherent inferences arise from relevance strategies with high values of intellectual obligation. In any event, the analyst reviews the coherence of the five relevance strategy-specific judgments in the cells of a column and enters a composite judgment in the "column summary" cell for each claim in Working Table 3.

Coherence then must be examined across the claims in the "Column Summary" row. The intent here is to determine whether the separate claims present a consistent pattern. For example, a claim that a substance is not genotoxic might be judged inconsistent with a claim that the substance induces neoplastic conversion (if it is premised that conversion arises from genotoxicity). This consideration might cause the analyst to adjust the claim to neoplastic conversion. This second level of summary, taking into consideration coherence across the columns of the "Column Summary", is provided in the "Overall Summary" of Working Table 3. The summary row in Working Tables 6 and 7 have the same function as in Working Table 3. In Working Table 5 the single summary row labeled "Overall Assessment" is like "Column Summary" in Working Table 3. There is no need for an additonal summary row in Working Table 5 because coherence of column summaries is not an issue.

Working Table 7 examines the coherence of Target-Context claims (the product of one of the forms of Working Table 3 as generated for the specific target context) and the inter-context extrapolation claims from any observational context to the target context (the product of one of the forms of Working Table 6). The analyst simply transcribes the results of the "Overall Summary" rows in Working Tables 3 and 6. The analysis of coherence for Working Table 7 then proceeds as in the discussion of Working Table 3. Coherence across contexts first is judged within a column (the "Column Summary" of Working Table 7), followed by a judgment of coherence across the claims in the "Column Summary" row. The result is the "Overall Summary" row of Working Table 7.

6. THE SEVEN STEPS OF HAZARD IDENTIFICATION: AN OVERVIEW

The seven steps indicated for hazard identification of a substance are shown schematically in Figures 4 and 5. The elements within each step are described verbally in Figure 4. In Figure 5 the steps are described in terms of the working tables to be completed and the inter-relationship between steps is depicted. The following brief discussion of the seven steps supplements the descriptions given

in the two figures. References are made to preceding sections of the report for discussion of topics relevant to each step.

Steps 1 and 2 refer to the initial collection of all informational sources that may be potential useful in assessing the support for claims of carcinogenicity of the substance of interest and definition of contexts (Section 4.1). "Observational contexts" are defined around studies of dose-response data in animals or humans while "target contexts" are defined for contexts of interest for hazard identification in which no study data are available. Studies judged to be sufficiently similar to justify pooling their results statistically are treated as "cases" belonging to the same observational context.

Step 3 is repeated for each observational context. Working Tables 2 and 3 (WT2 and WT3) are completed for each context. The data items in Table 2 for which data are available are inserted in WT2 (Step 3.2) and assessments of the data characteristics (completeness, utility, observed effect, and causality) are made (Step 3.3) (Sec. 4.2 and subsections). For contexts with more than one case, a similar step must first be conducted for each case within the context to form a composite representation of the context (Step 3.1). The completed WT2 for an observational context is utilized in completion of WT3 for the same context, in which a judgment of support for each claim of carcinogenicity (c. of c.) is made for each relevance strategy available (Step 3.4). The upper part of an entry in WT3 is for a judgment based on data used only in that context (e.g., tumor response data items); the lower part of the entry is for a judgment based on all data relevant to a judgment of carcinogenicity within the context, including data that may be utilized in other contexts as well (e.g., biophysical effects or pharmacodynamic effects obtained *in vitro*, structure activity relationships, etc.) (Sec. 5.2). After completion of the entries of WT3 for an observational context, the support for each claim of carcinogenicity is summarized (from the lower part of entries) in the "column summary" of WT3 (Step 3.4). Upon reflection of the assessments shown in the column summary for coherence, the "overall summary" entries are completed in WT3 (Step 3.5).

Figure 4. The Seven Steps of Hazard Identification in Carcinogen Risk Analysis

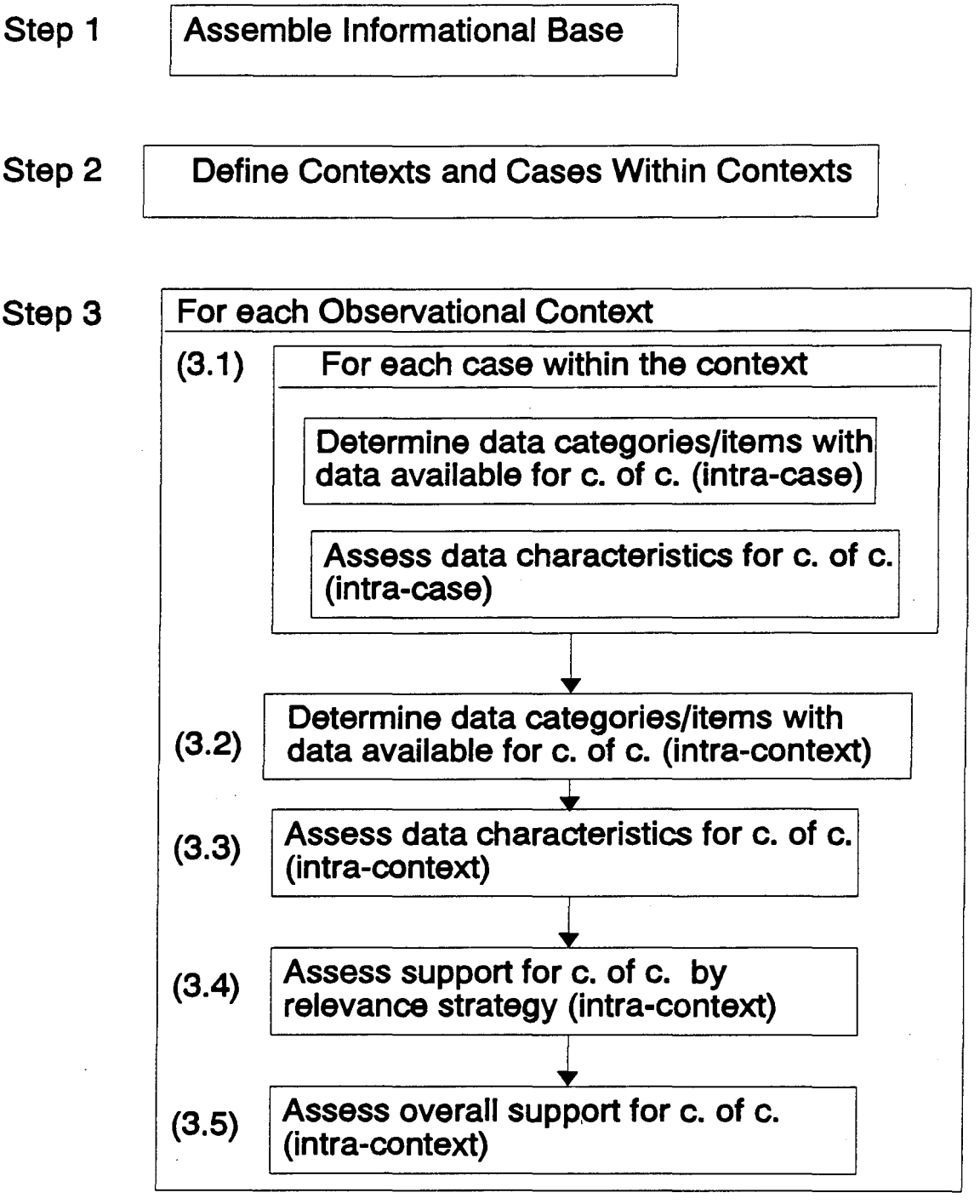


Figure 4. (continued)

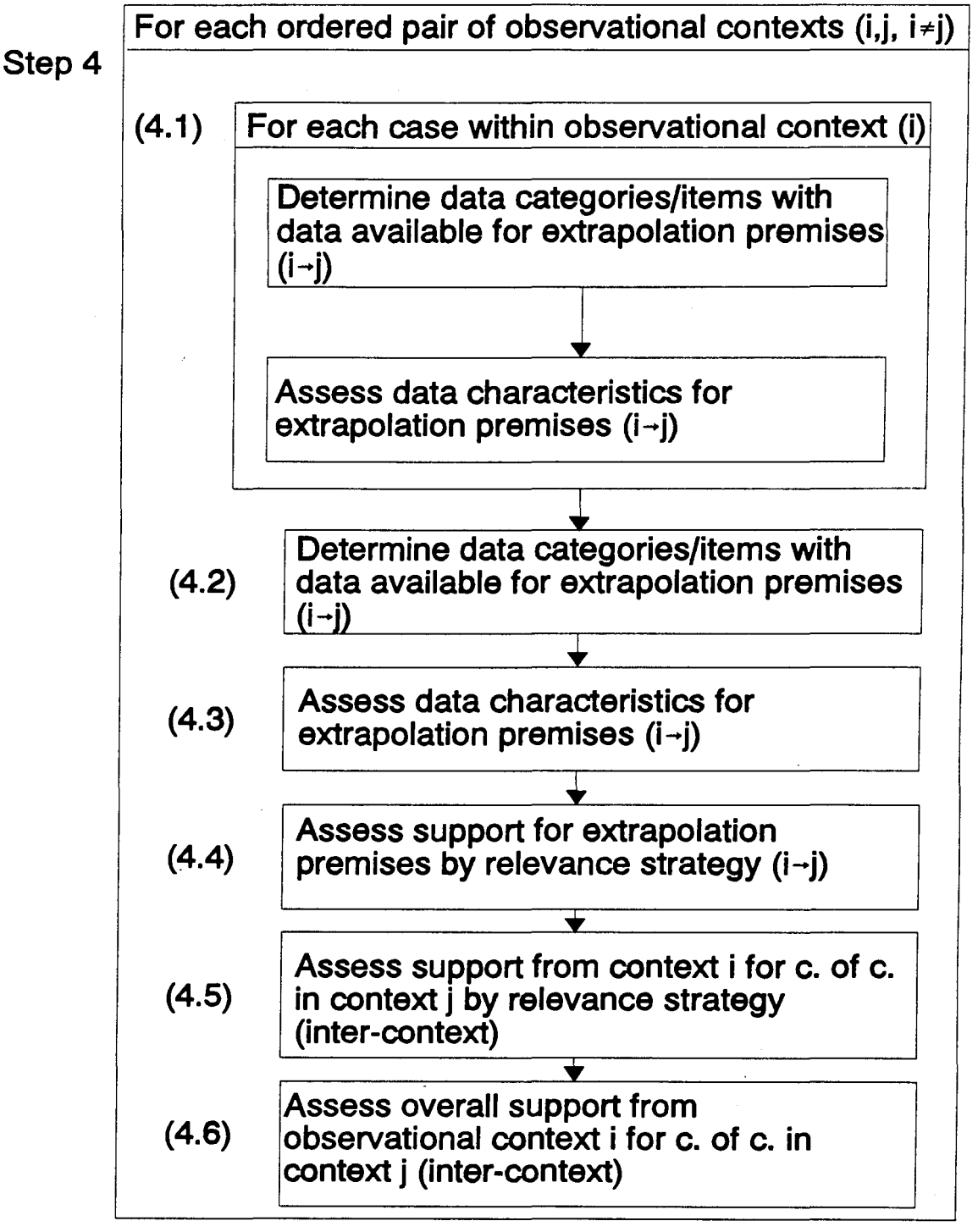


Figure 4. (continued)

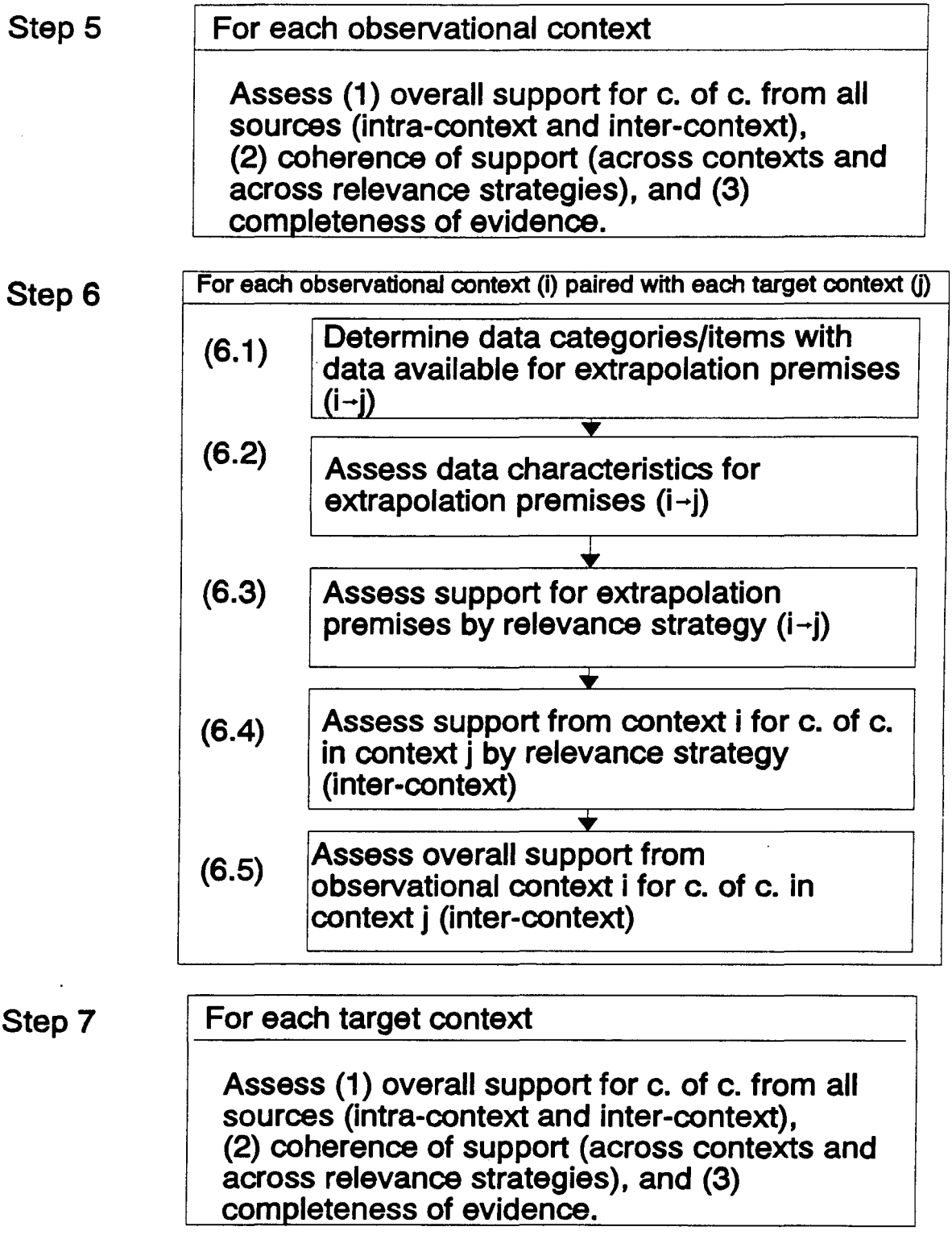


Figure 5. Flow Diagram for the Seven Steps of Hazard Identification with Application of Working Tables

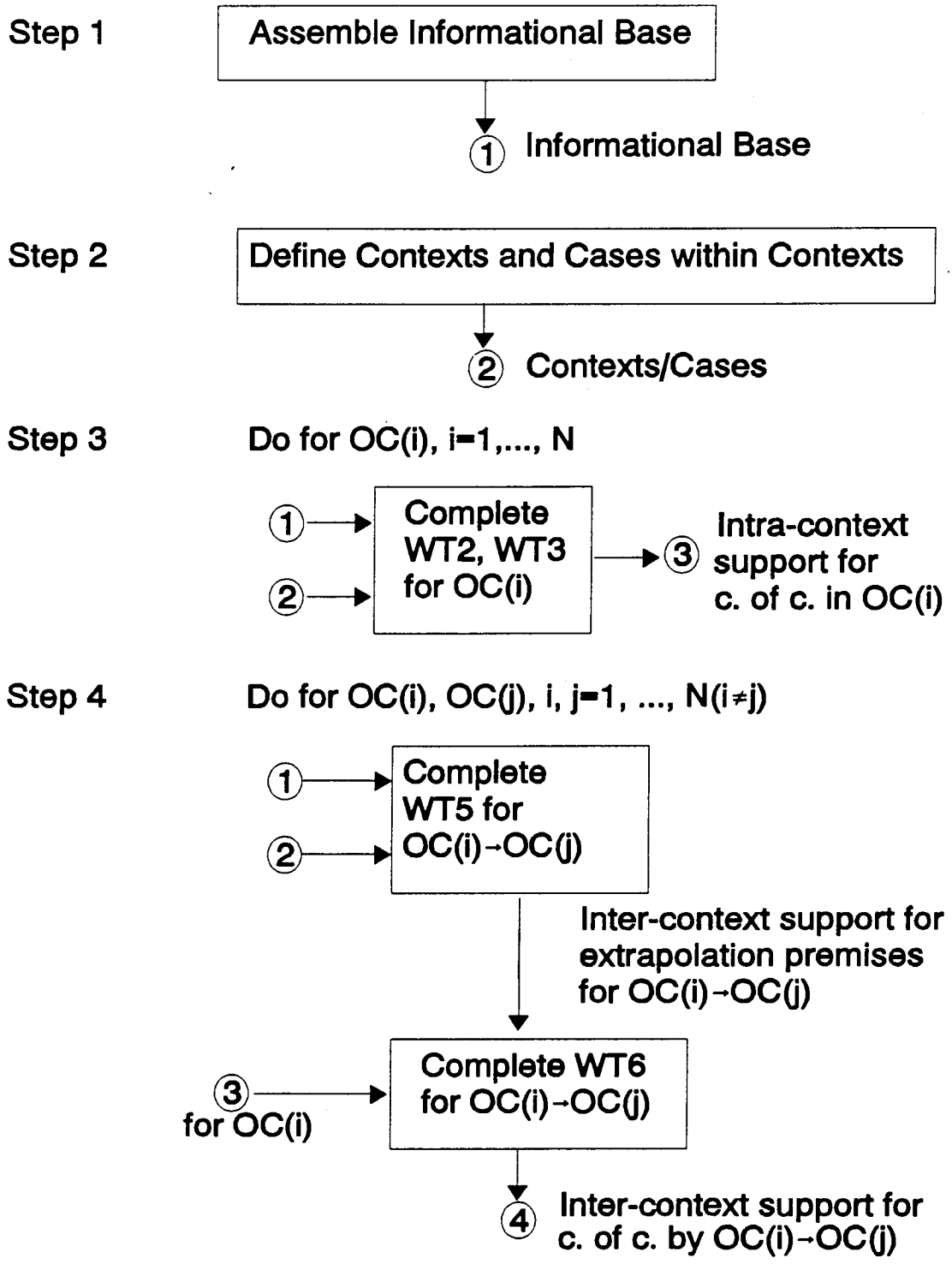
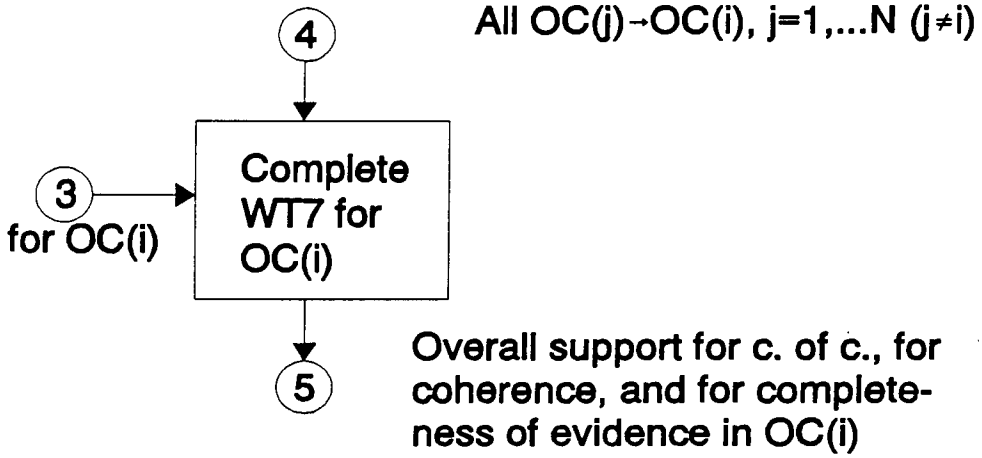


Figure 5. (continued)

Step 5 Do for OC(i), i=1,...N



Step 6 Do for OC(i), TC(j), i=1,..., N, j=1,...M

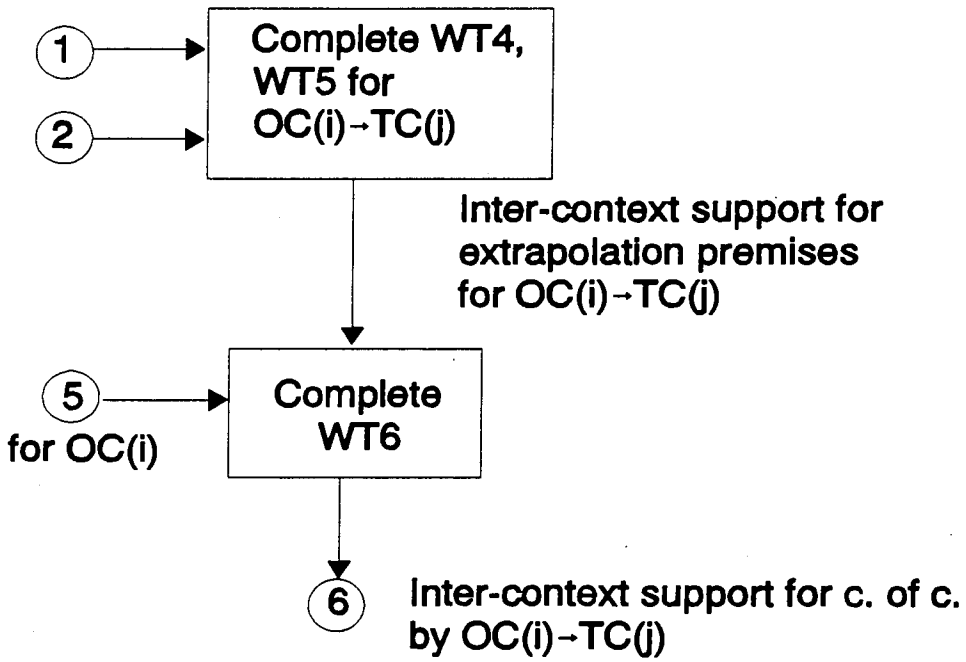
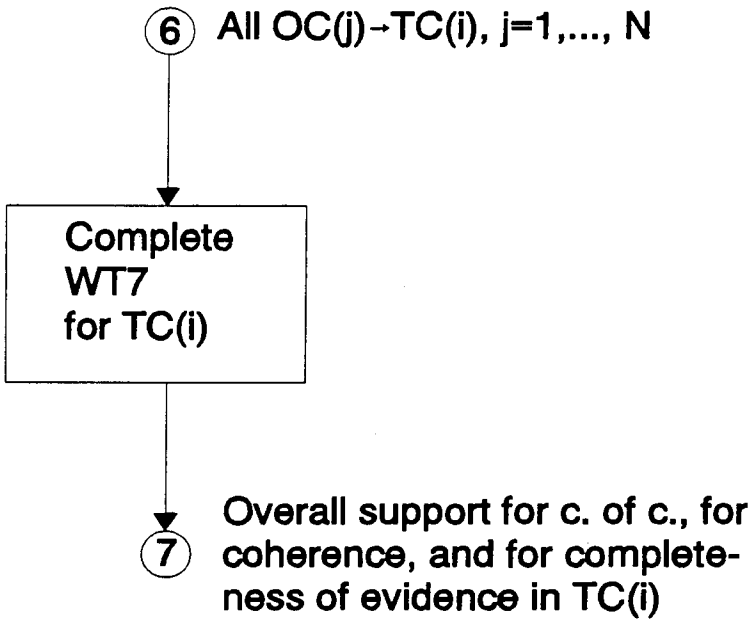


Figure 5. (continued)

Step 7 Do for TC(i), i=1,..., M



Key to abbreviations:

c. of c.: claims of carcinogenicity
OC(i): observational context i
TC(i): target context i
①,②,,,: output from Steps 1, 2,...
OC(i)→OC(j): "Extrapolation from observational context i to observational context j"
OC(i)→TC(j): "Extrapolation from observational context i to target context j"
N: Number of observational contexts
M: Number of target contexts
WT1: Working Table 1 (similarly for WT2, WT3,...)

Step 4 deals with extrapolating claims of carcinogenicity from one observational context (e.g., the jth observational context, denoted as OC(i) in Figure 4) to another observational context (e.g., the jth observational context, denoted as OC(j)). The necessity for this step arises from the need to base the judgment for claims of carcinogenicity in each observational context on all the information that may contain some evidential basis. At this point we have an intra-context assessment for each observational context, i.e., only the data from within the context itself has been utilized. In Step 4, Working Tables 4, 5, and 6 are completed for each pair of observational contexts. This step for inter-context support of carcinogenicity bears some similarity to Step 3 for intra-context support. Instead of completing WT2 for data related to carcinogenicity within the context, however, WT4 (with the same headings) is completed for data items needed to determine the support for extrapolating the intra-context claims for carcinogenicity in OC(i) to OC(j). The data items for this purpose are listed in Table 2 and the role of each data category is described in Table 9. Steps 4.1-4.4 are conceptually analogous to Steps 3.1-3.4 with the difference being that the endpoint is a judgment of support for extrapolation premises from OC(i) to OC(j), instead of intra-context support for carcinogenicity (Sec. 5.3). Continuing the analogy, WT4 and WT5 play the previous roles of WT2 and WT3. The overall objective of Step 4, however, is a judgment of support for claims of carcinogenicity in OC(j) based on extrapolation from OC(i). That assessment is made in Step 4.5 (which has no counterpart in Step 3), wherein entries in WT6 are completed based on the support for extrapolation premises (in WT5) and the intra-context support for carcinogenicity in OC(i) (in WT3). It may be noted that WT6 is identical in format to WT3, aside from the heading. The remaining portions of WT6 are completed in the same way as WT3.

Step 5 summarizes the overall support for each observational context from the intra-context assessment for the context (WT3) and the inter-context assessments from all other observational

contexts (WT6), which are entered into WT7. It remains to assess support for claims of carcinogenicity in target contexts.

Steps 6 and 7 accomplish the assessment of support for claims of carcinogenicity in target contexts. Since there is no intra-context assessment (because there are no observational data, by definition), judgments are based totally on extrapolation from observational contexts. For each observational context, data items related to premises for extrapolation to the target context of interest are assembled and evaluated in the same manner as in extrapolation between observational contexts. WT4 and WT5 are completed as in Step 4. The judgment of the support for claims of carcinogenicity in the target context from each observational context is based on the support for extrapolation premises (in WT5) and the overall support for claims of carcinogenicity in the observational context (in WT7 for the observational context), and is entered into WT6 for the target context. The overall support for claims of carcinogenicity in the target context is based on the support determined by extrapolation from each observational context, and is entered in WT7 for the target context.

REFERENCES (PART II)

1. Aaronson, A.; Tronick, S. (1986) The Role of Oncogenes in Human Neoplasia. *Important Advances in Oncology*. Philadelphia: Lippincott Publishers.

2. Alarie, Y. (1981) Dose-Response Analysis in Animal Studies: Prediction of Human Responses. *Environmental Health Perspectives* 42:9.

3. Allen, B.; Crump, K.; Shipp, A. (1988) Is it possible to predict the carcinogenic potency of a chemical in humans using animal data? *Banbury Report 31: Carcinogen Risk Assessment: New Directions in the Qualitative and Quantitative Aspects.*

4. Alston, W. (1985) Concepts of Epistemic Justification. *The Monist* 68:57-89.

5. Ames, B.; Gold, L. (1991a) Mitogenesis, Mutagenesis and Animal Cancer Tests. *Chemically Induced Cell Proliferation: Implications for Risk Assessment.* New York: Wiley and Sons, pp. 1-20.

6. Ames, B.M.; Gold, L.S. (1991b) Animal Cancer Tests and the Prevention of Cancer. Proceedings of the International Council for Coordinating Cancer Research Conference.

7. Ames, B.M.; Gold, L.S. (1990) Chemical carcinogenesis: Too many rodent carcinogens. *Proc Natl Acad Sci USA* 87:7772-7776.

8. Ames, B.M.; McCann, J.; Yamasaki, E. (1975) Method for detecting carcinogens and mutagens with the Salmonella/mammalian microsome test. *Mutation Research* 31:347-364.

9. Andersen, M. (1981) A Physiologically Based Toxicokinetic Description of the Metabolism of Inhaled Gases and Vapors: Analysis of Steady State. *Toxicology and Applied Pharmacology* 60:509.

10. Andersen, M.; Archer, R.; Clewell, H.; MacNaughton, M. (1984) A Physiological Model of the Intravenous and Inhalation Pharmacokinetics of Three Dihalomethanes--CH_2Cl_2, $CH_{2r}Cl$ and CH_2Br_2--in the Rat. *Toxicologist* 4:111.

11. Andersen, M.; Garags, M.; Clewell, M. (1986) Suicide Inactivation of Microsomal Oxidation by cis- and trans-dichloroethylene (C-DCE and T-DCE) in Male Fisher Rats In-vivo. *Toxicologist* 6:12.

12. Andersen, M.; Gargas, M.; Ramsey, J. (1987) Inhalation Pharmacokinetics: Evaluating Systemic Extraction, Total In-vivo Metabolism and the Time Course of Enzyme Induction for Inhaled Styrene in Rats Based on Arterial Blood: Inhaled Air Concentration Ratios. *Toxicology and Applied Pharmacology* 73:176.

13. Andersen, M.E. (1989) Tissue Dosimetry, Physiologically-Based Pharmacokinetic Modeling, and Cancer Risk Assessment. *Cell Biology and Toxicology* 5(4):405-415.

14. Andersen, M.E.; Clewell III, H.J.; Gargas, M.L.; Smith, F.A.; Reitz, R.H. (1987) Physiologically Based Pharmacokinetics and the Risk Assessment Process for Methylene Chloride. *Toxicology and Applied Pharmacology* 87:185-205.

15. Andersen, M.E.; Clewell III, H.J.; Reitz, R.H. (1987) Physiologically based pharmacokinetics and methylene chloride cancer risk assessment considerations. *Human Risk Assessment--The Role of Animal Selection and Extrapolation*. London: Taylor and Francis Publishers, pp. 129-139.

16. Anderson, M.W. (1987) Issues in Biochemical Applications to Risk Assessment: How Do We Evaluate Individual Components of Multistage Models? *Environmental Health Perspectives* 76:175-179.

17. Anderson, M.W.; Hoel, D.G.; Kaplan, N.L. (1980) A General Scheme for the Incorporation of Pharmacokinetics in Low-Dose Risk Estimation for Chemical Carcinogenesis: Example--Vinyl Chloride. *Toxicology and Applied Pharmacology* 55:154-161.

18. Archer, L.E.; Ryan, L.M. (1989) On the Role of Cause-of-Death Data in the Analysis of Rodent Tumorgenicity Experiments. *Appl Statist* 38(1):81-93.

19. Argus, M.; Hoch-Ligetic, C.; Arcos, J.; Conney, A. (1978) Differential Effects of ß-napthoflavone and pregnenolone-16-γ-carbonitrile on dimethylnitrosamine Induced Hepatocarcinogenesis. *Journal of the National Cancer Institute* 61:441.

20. Arms, A.D.; Travis, C.C. (1988) Reference Physiological Parameters in Pharmacokinetic Modeling. United States Environmental Protection Agency. EPA/600/6-88/004.

21. Ashby, J. (1986) Genetic Toxicology in Industry: Perspectives and Initiatives. *Risk and Reason: Risk Assessment in Relation to Environmental Mutagens and Carcinogens*. New York: Alan R. Liss, Inc., pp. 89-94.

22. Ashby, J. (1990) Genotoxicity Testing: To What Extent Can It Recognize Mutagens and Carcinogens? *Progress in Predictive Toxicology*. New York: Elsevier, pp. 185-205.

23. Ashby, J.; Morrod, R.S. (1991) Detection of Human Carcinogens. *Nature* 352:185-186.

24. Ashby, J.; Tennant, R.W. (1991) Definitive Relationships among Chemical Structure, Carcinogenicity and Mutagencity for 301 Chemicals Tested by the US NTP. *Mutation Research* 257:229-306.

25. Ashby, J.; Tennant, R.W. (1988) Chemical Structure, Salmonella Mutagenicity and Extent of Carcinogenicity as Indicators of Genotoxic Carcinogenesis among 222 Chemicals tested in Rodents by the US NTP. *Mutation Research* 204:17-115.

26. Ashby, J.; Tennant, R.W.; Zeiger, E.; Stasiewicz, S. (1989) Classification according to Chemical Structure, Mutagenicity to Salmonella and level of Carcinogenicity of a further 42 Chemicals Tested by the US NTP. *Mutation Research* 223:73-103.

27. Bailey, A.; Sanchez, J.; Rodu, B.; Shrestha, K.; Tran-Paterson, R.; Saliba, D.; Miller, D. (1991) Induced Proto-Oncogene Expression Increases Susceptibility to Radiation-Induced Carcinogenesis. *Chemically Induced Cell Proliferation: Close Up Implications for Risk Assessment*. New York: Wiley and Sons, pp. 105-114.

28. Bannasch, P.; Enzmann, H.; Zerban, H. (1987) Preneoplastic Lesions as Indicators of the Carcinogenic Risk Caused by Chemicals. *Cancer Risks: Strategies for Elimination*. Springer Verlag, Berlin, pp. 47-64.

29. Barrett, J. (1984) Tumor Promotion and Tumor Progression. *Carcinogenesis-A Comprehensive Survey*, vol. 8. New York: Raven Press, pp. 423-429.

30. Barrett, J. (1987) A Multistep Model for Neoplastic Development: Role of Genetic and Epigenetic changes. *Mechanisms of Environmental Carcinogenesis*, Vol. II. Florida: CRC Press, pp. 117-126.

31. Barrett, J.C.; Kakunaga, T.; Kuroki, T.; Neubert, D.; Trosko, J.E.; Vasiliev, J.M.; Williams, G.M.; Yamasaki, H. (1986a) Report 10: In-Vitro Assays That May Be Predictive Of Tumour-Promoting Agents. *Long-Term and Short-Term Assays for Carcinogens: A Critical Appraisal: IARC* 83:287-302.

32. Barrett, J.C.; Kakunaga, T.; Kuroki, T.; Neubert, D.; Trosko, J.E.; Vasiliev, J.M.; Williams, G.M.; Yamasaki, H. (1986b) Report 9: Mammalian Cell Transformation in Culture. *Long-Term and Short-Term Assays for Carcinogens: A Critical Appraisal: IARC* 83:267-286.

33. Barrett, J.C.; Wiseman, R.W. (1987) Cellular and Molecular Mechanisms of Multistep Carcinogenesis: Relevance to Carcinogen Risk Assessment. *Environmental Health Perspectives* 76:65-70.

34. Bartsch, H.; Armstrong, B. (eds.) (1982) *Host Factors in Human Carcinogenesis*.

35. Bartsch, H.; Malaveille, C. (1990) Screening assays for carcinogenic agents and mixtures: an appraisal based on data in the IARC Monograph series. *Complex Mixtures and Cancer Risk*, pp.65-74.

36. Baty, D. (1979) Species, Strain and Sex Differences in Metabolism. *Foreign Comp Metabolism in Mammals* 5:159.

37. Baty, D. (1981) Species, Strain and Sex Differences in Metabolism. *Foreign Comp Metabolism in Mammals* 6:133.

38. Beaune, P.; Kremers, P.; Kaminsky, L.; DeGraeve, J.; Guengerich, F. (1986) Comparison of Monooxygenase Activities and Cytochrome P-450 Isozyme Concentrations in Human Liver Microsomes. *Drug Metabolism and Disposition: The Biological Fate of Chemicals* 14:437.

39. Belinsky, S.; Walker, V.; Maronpot, R.; Swenberg, J.; Andersen, M. (1987a) Molecular Dosimetry of DNA Adduct Formation and Cell Toxicity in Nasal Mucosa Following Exposure to a Tobacco Specific Nitrosamine: Relationship to Induction of Neoplasia. *Cancer Research* 47:6058.

40. Belinsky, S.; White, C.; Devereux, T.; Swenberg, J.; Andersen, M. (1987b) Cell Selective Alkylation of DNA in Rat Lung Following Low Dose Exposure to the Tobacco Specific Carcinogen 4-(n-methyl-n-nitrosamine)-1-(3-pyridyl)-1-butanone. *Cancer Research* 47:1143.

41. Belinsky, S.A.; White, C.M.; Devereux, T.R.; Anderson, M.W. (1987c) DNA Adducts as a Dosimeter for Risk Estimation. *Environmental Health Perspectives* 76:3-8.

42. Bernblum, I. (1985) Challenging Problems in Cocarcinogenesis. *Cancer Research* 45:1917-1921.

43. Berenblum, I.; Haran-Ghera, N. (1957) A Quantitative Study of the Systemic Initiating Action of Urethane (Ethyl Carbamate) in Mouse Skin Carcinogenesis. *British Journal of Cancer* 11:77-84.

44. Bernstein, R. (1983) *Beyond Objectivism and Relativism*. Philadelphia, PA: University of North Carolina Press.

45. Bertram, J. (1990) Role of Gap Junctional Cell/Cell Communication in the Control of Proliferation and Neoplastic Transformation. *Radiation Research* 123:252-256.

46. Bickis, M.; Krewski, D. (1985) Statistical Design and Analysis of the Long-Term Carcinogenicity Bioassay. *Toxicological Risk Assessment, Vol. 1: Biological and Statistical Criteria*. Florida: CRC Press, pp. 125-147.

47. Bickis, M.; Krewski, D. (1989) Statistical Issues in the Analysis of the Long-Term Carcinogenicity Bioassay in Small Rodents: An Empirical Evaluation of Statistical Decision Rules. *Fundamental and Applied Toxicology* 12:202-221.

48. Boberg, E.; Miller, E.; Miller, J.; Poland, A.; Liem, A. (1983) Strong Evidence from Studies with Brachymorphic Mice and Pentachlorphenol That 1'-sulfooxysafrole is the Major Electrophilic and Carcinogenic Metabolite of 1'-hydroxysafrole in Mouse Liver. *Cancer Research* 43:5163.

49. Bogen, K. (1989) Cell Proliferation Kinetics and Multi-Stage Cancer Risk Models. *Journal of National Cancer Institute* 81:267-277.

50. Boorman, G.A.; Montgomery Jr., C.A.; Eustis, S.L.; Wolfe, M.J.; McConnell, E.E.; Hardisty, J.F. (1985) Chapter 16. Quality Assurance in Pathology for Rodent Carcinogenicity Studies. *Handbook of Carcinogen Testing*. New Jersey: Noyes Publications.

51. Bos, J. (1988) The ras Gene Family and Human Cancer. *Mutation Research* 195:255-271.

52. Boutwell, R. (1974) The Function and Mechanism of Promoters of Carcinogenesis. *CRC Critical Reviews of Toxicology* 2:419-443.

53. Brandom, W.; Scaaomanno, G.; Archer, V.; Arhcer, P.; Bloom, A. (1978) Chromosome Aberrations as a Biological Dose-Response Indicator of Radiation Exposure in Uranium Miners. *Radiation Research* 76:159-171.

54. Breslow, N.E.; Day, N.E. (1980) *Statistical Methods in Cancer Research. Volume I. The Analysis of Case-Control Studies.*

55. Breslow, N.E.; Day, N.E. (1987) *Statistical Methods in Cancer Research. Volume II. The Design and Analysis of Cohort Studies.*

56. Brown, I. (1988) *Rationality.* Routledge, London.

57. Burchell, B.; Jackson, M.; Fennedy, S.; McCarthy, L.; Barr, G. (1985) Characterization and Regulation of Hepatic UDP-glucuronyltransferase. *Microsomes and Drug Oxidations.* A. Boobis, et al., ed.; pp. 212-220.

58. Burke, M.; Upshall, D. (1976) Species and Phenobarbitone-induced Differences in the Kinetic Constants of Liver Microsomal Harmine O-demethylation. *Xenobiotica* 6:321.

59. Burkhart, K.; Kulig, K.; McMartin, K. (1990) Formate Levels Following a Formalin Ingestion. *Vet. Human Toxicology* 3:157-184.

60. Butterworth, B.E. (1990) Consideration of both Genotoxic and Nongenotoxic Mechanisms in Predicting Carcinogenic Potential. *Mutation Research* 239:117-132.

61. Butterworth, B.; Slaga, T.; Farland, W.; McClaim, M. (1989) *Chemically Induced Cell Proliferation: Implications for Risk Assessment.* Progress in Clinical and Biological Research, vol. 369.

62. Casanova-Schmitz, M.; Heck, H. (1983) Effects of Formaldehyde Exposure on the Extractability of DNA from Proteins in the Rat Nasal Mucosa. *Toxicology and Applied Pharmacology* 70:121.

63. Casanova-Schmitz, M.; Starr, T.; Heck, H. (1984) Differentiation Between Metabolic Incorporation and Covalent Binding in the Labeling of Macromolecules in the Rat Nasal Mucosa and Bone Marrow by Inhaled C^{14} and H^{3} Formaldehyde. *Toxicology and Applied Pharmacology* 76:26.

64. Chin, T.; Heck, H. (1981) Disposition of C^{14} Formaldehyde in Fisher-344 Rats Following Inhalation Exposure. *Toxicologist* 1:78.

65. Clayson, D.B. (1985) Dose Relationships in Experimental Carcinogenesis: Dependence on Multiple Factors Including Biotransformation. *Toxicologic Pathology ISSN: 0192-6233* 13(2):119-127.

66. Clayson, D.B. (1987) ICPEMC Publication No. 13: The Need for Biological Risk Assessment in Reaching Decisions about Carcinogens. *Mutation Research* 185:243-369.

67. Clayson, D.B. (1988) Needs for Biological Risk Assessment in Interspecies Extrapolation. *Environmental Health Perspectives* 77:93-97.

68. Clayson, D.B.; Arnold, D.L. (1991) ICPEMC Publication No. 19: The Classification of Carcinogens Identified in the Rodent Bioassay as Potential Risks to Humans: What Type of Substance Should be Tested Next? *Mutation Research* 257:91-106.

69. Clayson, D.B.; Krewski, D. (1986) ICPEMC Working Paper No. 3: The Concept of Negativity in Experimental Carcinogenesis. *Mutation Research* 167:233-40.

70. Clayson, D.B.; Krewski, D.; Munro, I.C. (1983) The Power and Interpretation of the Carcinogenicity Bioassay. *Regulatory Toxicology and Pharmacology* 3:329-348.

71. Clewell III, H.J.; Andersen, M.E. (1989) Biologically Motivated Models for Chemical Risk Assessment. *Health Physics* 57(1):129-137.

72. Code of Federal Regulations. (1983a) Toxic Substances Control: Good Laboratory Practice Standards-Final Rule. C.F.R. §40:53922-53944.

73. Code of Federal Regulations. (1983b) Pesticide Programs: Good Laboratory Practice Standards-Final Rule. C.F.R. §40:53946-53969.

74. Cohen, S.M.; Ellwein, L.B. (1990) Cell Proliferation in Carcinogenesis. *Science* 249:1007-1011.

75. Cohrssen, J.J.; Covello, V.T. (1989) *Risk Analysis: A Guide to Principles and Methods for Analyzing Health and Environmental Risks.*

76. Cone, J.E.; Reeve, G.R.; Landrigan, P.J. (1987) Chapter 6. Clinical and Epidemiological Studies. *Toxic Substances and Human Risk.* New York: Plenum Press.

77. Conney, A. (1982) Induction of Microsomal Enzymes by Foreign Chemicals and Carcinogenesis by Polycyclic Aromatic Chemicals. *Cancer Research* 82:4975.

78. Conney, A.; Buening, M.; Pantuck, E.; Pantuck, C.; Fortner, J.; Anderson, K.; Kappas, A. (1980) Regulation of Human Drug Metabolism by Dietary Factors. *Environmental Chemicals, Enzyme Function and Human Disease*, Ciba Foundation Symposium Vol. 76.

79. Cook, R.R. (1982) The Role of Epidemiology in Risk Assessment. *Drug Metabolism Reviews* 13(5):913-923.

80. Cooper, G.M. (1990) Chapter 1: The Cancer Cell. *Oncogenes.* Boston: Jones and Bartlett Publishers, pp. 3-18.

81. Counts, C; Bruckner, J.; Feldman, S. (1982) Role of Dose Level, Food Intake and Diluent in Toxicokinetics of Orally Administered 1,1-dichloroethylene (1,1 DCE). *Pharmacologist* 24:171.

82. Craig, P.; Enstein, K. (1981) Structure-Activity in Hazard Assessment. *Hazard Assessment of Chemicals: Current Development, Vol. 1.* New York: National Academy Press, pp. 389-420.

83. Crawford-Brown, D.; Hofmann, W. (1989) The Role of Variability of Dose in Dose-Response Relationships for Alpha Emitting Radionuclides. *Radiation Protection Dosimetry* 28:283-290.

84. Crawford-Brown, D.; Hofmann, W. (1990) A Generalized State-Vector Model for Radiation-Induced Cellular Transformation. *International Journal of Radiation Biology* 57:407-423.

85. Cross, F. (1987) Health Effects. *Envrionmental Radon.* New York: Plenum Press, pp. 215-243.

86. Crump, K.; Allen, B. (1984) Quantitative Estimates of Risk from Occupational Exposure to Benzene. Report to OSHA in May.

87. Curtis, S. (1991) Mechanisitic Models. *Physical and Chemical Mechanisms in Molecular Radiation Biology.* New York: Plenum Press, pp. 367-385.

88. Dahlman, T.; Rhodin, J. (1956) Effects of Exposure to Lung-Irritating Gas on Transport of Secretion and Fibrillation Frequency in Rat Trachea. *Nordisk Hygienisk Tidskrist* 37:82.

89. Day, N.E. (1988) Chapter 1. Epidemiological Methods for the Assessment of Human Cancer Risk. *Toxicological Risk Assessment, Vol. II.* Florida: CRC Press.

90. Della-Favera, R.; Martinotti, S.; Gallo, R.; Erickson, J.; Croce, C. (1988) Translocation and Rearrangements of the c-myc Oncogene Locus in Human Undifferentiated B-cell Lymphomas. *Science* 219:963-997.

91. DerSimonian, R.; Laird, N. (1986) Meta-analysis in Clincal Trials. *Controlled Clinical Trials* 7:177.

92. Dietz, F.; Ramsey, J.; Watanabe, P. (1983) Relevance of Experimental Studies to Human Risk. *Environmental Health Perspectives* 52:9.

93. Doll, R. (1985) General Conclusions. *Interpretation of Negative Epidemiological Evidence for Carcinogenicity, IARC Scientific Publications No. 65.*

94. Dybing, E. (1986) Predictability of Human Carcinogenicity from Animal Studies. *Regulatory Toxicology and Pharmacology* 6:399.

95. Eccles, S.A. (1987) Chapter 6: Animal models of cancer. *Biology of Carcinogenesis.* Lancaster: MTP Press Limited, pp.109-139.

96. Eddy, D.M. (1989) The Confidence Profile Method: A Bayesian Method for Assessing Health Technologies. *Oper Res* 37:210.

97. Eddy, D.M.; Hasselblad, V.; Shachter, R. (1990a) An Introduction to a Bayesian Method for Meta-analysis: The Confidence Profile Method. *Med Decision Making* 10:15.

98. Eddy, D.M.; Hasselblad, V.; Shachter, R. (1990b) A Bayesian Method for Synthesizing Evidence: The Confidence Profile Method. *Int. J. Technol. Asses. Health Care* 6:31.

99. Environmental Health Criteria 109 (1990) *Summary Report on the Evaluation of Short-term Tests for Carcinogenesis (Collaborative Study on In-vivo Tests)*. World Health Organization.

100. European Chemical Industry Ecology and Toxicology Center (1981) The Mutagenic and Carcinogenic Potential of Formaldehyde Technical Report No. 2. *ECETOC*, Brussels.

101. Evans, A. (1978) Casusation and Disease: A Chronological Journal. *American Journal of Epidemiology* 108:249-258.

102. Farber, E. (1987) Possible Etiologic Mechanisms in Chemical Carcinogenesis. *Environmental Health Perspectives* 75:65-70.

103. Farrar, D.B.; Crump, K.S. (1988) Exact Statistical Tests for Any Carcinogenic Effect in Animal Bioassays. *Fundamental and Applied Toxicology* 11:652-663.

104. Fiserova-Bergerova, V. (1976) Mathematical Modeling of Inhalation Exposure. *Journal of Combustion Toxicology* 3:201.

105. Fiserova-Bergerova, V.; Vlach, V.; Cassady, J. (1980) Predictable Individual Differences in Uptake and Excretion of Gases and Lipid Soluble Vapors: Simulation Study. *British Journal of Industrial Medicine* 37:42.

106. Friedman, M.; Wehr, C.; Schade, J.; MacGregor, J. (1982) Inactivation of Aflatoxin B_1 Mutagenicity By Thiols. *Food and Chemical Toxicology* 20:887.

107. Gargas, M.; Clewell, H.; Andersen (1986) Metabolism of Inhaled Dihalomethanes In-vivo: Differentation of Kinetic Constants for Two Independent Pathways. *Toxicology and Applied Pharmacology* 82:211.

108. Gart, J.J.; Chu, K.C.; Tarone, R.E. (1979) Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity. *Journal of the National Cancer Institute* 62(4):957.

109. Gart, J.J.; Krewski, D.; Lee, P.N.; Tarone, R.E.; Wahrendorf, J. (1986) *Statistical Methods in Cancer Research. Volume III. The Design and Analysis of Long-Term Animal Experiments.*

110. Gehring, P.; Watanabe, P.; Park, C. (1978) Resolution of Dose-Response Toxicity Data for Chemicals Requiring Metabolic Activation--Example-Vinyl Chloride. *Toxicology and Applied Pharmacology* 44:581.

111. Gelboin, H.; PTsO (1978) *Polycylic Hydrocarbons and Cancer, Vols. 1 and 2.* New York: Academic Press.

112. Gerlowski, L.; Jain, R. (1983) Physiologically Based Pharmacokinetic Modeling: Principles and Applications. *Journal of Pharmaceutical Science* 72:1103.

113. Ghoshal, A.; Rushmore, T.; Farber, E. (1987) Induction of Cancer by Dietary Deficiency without Added Carcinogens. *Nongenotoxic Mechanisms in Carcinogenesis: Banbury Report* 25:179-187.

114. Gil, J. (1982) Comparative Morphology and Ultrastructure of the Airways. *Mechanisms in Respiratory Toxicology.* Boca Raton, LA: CRC Press, pp. 3-25.

115. Gillette, J. (1976) Application of Pharmacokinetic Principles in the Extrapolation of Animal Data to Humans. *Clinical Toxicology* 9:709.

116. Gillette, J. (1984) Problems in Correlating In-vitro and In-vivo Studies of Drug Metabolism. *Pharmacokinetics: A Modern View.* New York: Plenum Press.

117. Graham, J.D.; Green, L.C.; Roberts, M.J. (1988) *In Search of Safety: Chemicals and Cancer Risk.* Massachusetts: Harvard University Press.

118. Greenland, S. (1987) Quantitative Methods in the Review of Epidemiologic Literature. *Epidemiologic Reviews* 9:1-30.

119. Gregory, A.R. (1988) Species Comparisons in Evaluating Carcinogenicity in Humans. *Regulatory Toxicology and Pharmacology* 8:160-190.

120. Gustafsson, J.; MacGeoch, C.; Morgan, E. (1985) Isolation Characterization and Regulation of Sex Specific Isozymes of Cytochrome P-450 Catalyzing 15 Beta-hydroxylation of Steroid Sulfates and 16-alpha-hydroxylation of 4- and rostene-3, 17-dione. *Microsomes and Drug Oxidations.* New York: Taylor & Francis.

121. Guyton, A. (1947) Measurement of the Respiration Volumes of Laboratory Animals. *American Journal of Physiology* 150:70.

122. Hamm Jr., T.E. (1985) Chapter 12. Design of a Long-Term Animal Bioassay for Carcinogenicity. *Handbook of Carcinogen Testing.* New Jersey: Noyes Publications, pp. 252-267.

123. Hall, E.; Freyer, G. (1991) The Molecular Biology of Radiation Carcinogenesis. *Physical and Chemical Mechanisms in Molecular Radiation Biology.* New York: Plenum Press, pp. 3-28.

124. Harper, B.L.; Legator, M.S. (1987) Tumor Promoters and Genotoxic Chemicals in Short-Term Testing for Carcinogenicity. *Advances in Veterinary Science and Comparative Medicine* 8:1-19.

125. Hartman, T.; Rosen, J. (1983) Inhibition of Metabolic Cooperation by Cigarette Smoke Condensate and Its Fractions in V-79 Chinese Hamster Lung Fibroblasts. Proceedings of the National Academy of Science 80:5305-5309.

126. Haseman, J.K. (1983) Statistical Support of the Proposed National Toxicology Program Protocol. *Toxicologic Pathology* 1(1):77.

127. Haseman, J.K. (1984) Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies. *Environmental Health Perspectives* 58:385-392.

128. Haseman, J.K. (1985) Issues in Carcinogenicity Testing: Dose Selction. *Fundamental and Applied Toxicology* 5:66-78.

129. Haseman, J.K. (1989) Commentary on "Are There Location/Cage/Systematic Nontreatment Effects in Long-Term Rodent Studies? A Question Revisited". *Fundamental and Applied Toxicology* 13:189-192.

130. Haseman, J.K. (1990) Use of Statistical Decision Rules for Evaluating Laboratory Animal Carcinogenicity Studies. *Fundamental and Applied Toxicology* 14:637-648.

131. Haseman, J.K.; Clark, A. (1990) Carcinogenicity Results for 114 Laboratory Animal Studies Used to Assess the Predictivity of Four In Vitro Genetic Toxicity Assays for Rodent Carcinogenicity. *Environmental and Molecular Mutagenesis* 16(18):15-31.

132. Haseman, J.K.; Huff, J.E. (1987) Species Correlation in Long-Term Carcinogenicity Studies. *Cancer Lettets* 37:125-132.

133. Haseman, J.K,; Huff, J.; Boorman, G.A. (1984) Use of Historical Control Data in Carcinogenicity Studies in Rodents. *Toxicologic Pathology* 12(2):126-135.

134. Haseman, J.K.; Huff, J.E.; Rao, G.N.; Eustis, S.L. (1989) Sources of Variability in Rodent Carcinogenicity Studies. *Fundamental and Applied Toxicology* 12:793-804.

135. Haseman, J.K.; Huff, J.E.; Zeiger, E.; McConnell, E.E. (1987) Comparative Results of 327 Chemical Carcinogenicity Studies. *Environmental Health Perspectives* 74:229-235.

136. Hattis, D. (1990) Pharmacokinetic Principles for Dose-Rate Extrapolation of Carcinogenic Risk from Genetically Active Agents. *Risk Analysis* 10(2):303.

137. Hattis, D.; White, P.; Marmorstein, L.; Koch, P. (1990) Uncertainties in Pharmacokinetic Modeling for Perchloroethylene. I. Comparison of Model Structure, Parameters, and Predictions for Low-Dose Metabolism Rates for Models Derived by Different Authors. *Risk Analysis* 10(3):449.

138. Hecker, E. (1984) Cocarcinogens of the Tumour-Promoter Type as Potential Risk Factors of Cancer in Man. *Models, Mechanisms and Etiology of Tumour Promotion: IARC* 56:441-463.

139. Hennings, H.; Shores, R.; Wenk, M.; Spangler, E.; Tarone, R.; Yuspa, S. (1983) Malignant Conversion of Mouse Skin Tumours is Increased by Tumour Initiators and Unaffected by Tumour Promoters. *Nature* 304:67-69.

140. Hermo Jr., H.; Brandt-Rauf, P.W. (1987) Chemically Induced Oncogenesis. *Seminars in Occupational Medicine* 2(4):235-43.

141. Heymann, E. (1980) Carboxylesterases and Amidase. *Enzymatic Basis of Detoxification, Vol. II.* New York: Academic Press.

142. Hill, A. (1965) The Environment and Disease: Association and Causation. *Proceedings of the Royal Society of Medicine* 58:295-300.

143. Himmelstein, K.; Lutz, R. (1979) A Review of the Application of Physiologically Based Pharmacokinetic Modeling. *Journal of Pharmacokinetics and Biopharmaceutics* 7:127.

144. Hoel, D.G. (1985) Epidemiology of Cancer Mechanisms. *National Cancer Inst Monogr* 67:199-203.

145. Hoel, D.; Kaplan, N.; Andersen, M. (1983) Implication of Nonlinear Kinetics on Risk Estimation in Carcinogenesis. *Science* 219:1032.

146. Hoel, D.G.; Haseman, J.K.; Hogan, M.D.; Huff, J.; McConnell, E.E. (1988) The Impact of Toxicity on Carcinogenicity Studies: Implications for Risk Assessment. *Carcinogenesis* 9(11):2045-2052.

147. Hoel, D.G.; Landrigan, P.J. (1987) Chapter 7. Comprehensive Evaluation of Human Data. *Toxic Substances and Human Risk.* New York: Plenum Press.

148. Hogben, C. (1971) Biological Membranes and Their Passage By Drugs. *Concepts in Biochemical Pharmacology,* Vol. 28:1-8.

149. Homburger, F.; Tregier, A. (1969) Modifiers of Carcinogenesis. *Progress in Experimental Tumor Research, Vol II.* New York: Karger, pp. 86-89.

150. Huff, J.E.; Eustis, S.L.; Haseman, J.K. (1989) Occurrence and Relevance of Chemically Induced Benign Neoplasms in Long-term Carcinogenicity Studies. *Cancer and Metastasis Reviews* 8:1-21.

151. Huff, J.E.; McConnell, E.E.; Haseman, J.K.; Boorman, G.A.; Eustis, S.L.; Schwetz, B.A.; Rao, G.N.; Jameson, C.W.; Hart, L.G.; Rall, D.P. (1986) Carcinogenesis Studies Results of 398 Experiments on 104 Chemicals. National Toxicology Program.

152. IARC, International Agency for Research on Cancer (1986) Long-Term and Short-Term Assays for Carcinogenesis, *IARC Scientific Publications,* No. 83.

153. Idle, J. (1980) Discussion. *Environmental Chemicals, Enzyme Function and Human Disease*, Ciba Foundation Symposium Vol. 76, p. 284.

154. Jakoby, W.; Sekura, R.; Lyon, E.; Marco, C.; Wang, J. (1980) Sulfotransferases. *Enzymatic Basis of Detoxification*. New York: Academic Press.

155. Kappas, A.; Alvarez, A.; Anderson, K.; Garland, W.; Pantuck, E.; Conney, A. (1977) The Regulation of Human Drug Metabolism by Nutritional Factors. *Drug Microsomes and Drug Oxidations*. New York: Pergamon Press.

156. Keck, G. (1981) Species Differences in Pharmacokinetics of Chemicals: Correlations with Toxicity. *Toxicology and Envrionmental Research* 3:207.

157. Kety, S. (1951) The Theory and Applications of the Exchange of Inert Gas at the Lungs and Tissues. *Pharmacological Reviews* 3:1.

158. Kinoshita, N.; Gelboin, G. (1972) Aryl Hydrocarbon Hydrolase and Polycyclic Hydrocarbon Tumorigenesis: Effect of the Enzyme Inhibitor 7,8 benzoflavone on Tumorigenesis and Macromolecular Binding. *Proceedings of the National Academy of Sciences* 69:824.

159. Kleinbloesem, C.; VanBrummeler, P.; Faber, H.; Danhof, M.; Vermuelen, N.; Breimer, D. (1984) Variability and Dynamics: A Net Oxidation Polymorphism in Man. *Biochemical Pharmacology* 33:3721.

160. Kopun, M.; Propping, P. (1977) The Kinetics of Ethanol Absorption and Elimination in Twins and Supplementary Repetitive Experiments in Singleton Subjects. *European Journal of Clinical Pharmacology* 11:337.

161. Krasovskii, G. (1976) Extrapolation of Experimental Data from Animals to Man. *Environmental Health Perspectives* 13:51.

162. Krewski, D.; Colin, D.; Dewanji, A. (1990) Chapter 16. Statistical Considerations in the Design of Toxicological Experiments with Chemical Mixtures. *Progress in Predictive Toxicology*. New York: Elsevier Publishers.

163. Krewski, D.; Goddard, M.J.; Murdoch, D. (1989) Statistical Considerations in the Interpretation of Negative Carcinogenicity Data. *Regulatory Toxicology and Pharmacology* 9:5-22.

164. Krewski, D.; Murdoch, D.J.; Withey, J.R. (1986) Paper: The Application of Pharmacokinetic Data in Carcinogenic Risk Assessment.

165. Krewski, D.; Smythe, R.T.; Dewanji, A. (1988) *Tests for Trend in Carcinogen Bioassay with Historical Controls*.

166. Krewski, D.; Wigle, D.; Clayson, D.B.; Howe, G.R. (1990) Role of Epidemiology in Health Risk Assessment. *Occupational Cancer Epidemiology*. New York: Springer-Verlad.

167. Kuhn, T. (1962) *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

168. Langenbach, R.; Tennant, R.W.; Barrett, J.C. (1988) Tumor Promotion: A Discussion of Current Topics, Available Assay Systems and Research Needs. *Tumor Promoters: Biological Approaches for Mechanistic Studies and Assay Systems*. New York: Raven Press, pp. 441-448.

169. Laudan, L. (1977) *Progress and Its Problems: Towards a Theory of Scientific Growth*. Berkley, CA: University of California Press.

170. Lawley, P.D. (1987) Concepts of Carcinogenesis. *Biology of Carcinogenesis*. Lancaster: MTP Press Limited, pp. 1-22.

171. Levin, W.; Thomas, P.; Reitz, L.; Ryan, D.; Bandiera, S.; Haniu, M.; Shivley, J. (1985) Immunochemical and Stuctural Characterization of Rat Hepatic Cytochrome P-450. *Microsomes and Drug Oxidations*. New York: Taylor & Francis.

172. Lewis, J.G.; Adams, D.O. (1987) Inflammation, Oxidative DNA Damage, and Carcinogenesis. *Environmental Health Perspectives* 76:19-27.

173. Longino, H. (1990) *Science as Social Knowledge*. Princeton, NJ: Princeton University Press.

174. Loury, D.J.; Goldsworthy, T.L.; Butterworth, B.E. (1987) The Value of Measuring Cell Replication as a Predictive Index of Tissue-specific Tumorigenic Potential. *Nongenotoxic Mechanisms in Carcinogenesis: Banbury Report* 25:119-136.

175. Lucier, G.W.; Thompson, C.L. (1987) Issues in Biochemical Applications to Risk Assessment: When Can Lymphocytes be used as Surrogate Markers? *Environmental Health Perspectives* 76:187-191.

176. Lutz, R.J.; Dedrick, R.L. (1987) Implications of Pharmacokinetic Modeling in Risk Assessment Analysis. *Environmental Health Perspectives* 76:97-106.

177. Maclure, M. (1985) Popperian Refutation in Epidemiology. *American Journal of Epidemiology* 121:343-350.

178. Mann, C. (1990) Meta-Analysis in the Breech. *Science* 249:476-480.

179. Mantel, N. (1980) Assessing Laboratory Evidence for Neoplastic Activity. *Biometrics* 36:381-399.

180. Maronpot, R.R. (1985) Chapter 18. Considerations in the Evaluation and Interpretation of Long-Term Animal Bioassays for Carcinogenicity. *Handbook of Carcinogen Testing*. New Jersey: Noyes Publications, pp. 372-383.

181. Mattison, D.R., Jelovsek, F.R. (1987) Pharmacokinetics and Expert Systems as Aids for Risk Assessment in Reproductive Toxicology. *Environmental Health Perspectives* 76:107-119.

182. McDougal, J.; Jepson, G.; Clewell, H.; MacNaughton, M.; Andersen, M. (1986) A Physiological Pharmacokinetic Model for Dermal Absorption. *Toxicology and Applied Pharmacology* 85:286.

183. McNight, B. (1988) A Guide to the Statistical Analysis of Long-Term Carcinogenicity Assays. *Fundamental and Applied Toxicology* 10:355-364.

184. Mettlin, C. (1987) Descriptive and Analytic Epidemiology. *Cancer* 62:1680-1687.

185. McBodile, M.; Holscher, M.; Neal, R. (1975) A Possible Protective Role for Reduced Glutathione in Aflatoxin B1 Toxicity: Effects of Pretreatment of Rats with Phenobarbital and 3-methylcholanthreneon Aflatoxin Toxicity. *Toxicology and Applied Pharmacology* 34:128.

186. Miller, E.; Miller, J. (1976) The Metabolism of Chemical Carcinogens to Reactive Electrophiles and Their Possible Mechanisms of Action in Carcinogenesis. *Chemical Carcinogens*. Washington, D.C.: American Chemical Society, p.737.

187. Miller, R.; Guengerich, F. (1983) Metabolism of Trichloroethylene in Isolated Hepatoxytes, Microsomes and Reconstituted Enzyme Systems Containing Purified Cytochromes P-450. *Cancer Research* 43:1145.

188. Moolgavkar, S. (1991) Carcinogenesis Models: An Overview. *Physical and Chemical Mechanisms in Molecular Radiation Biology*. New York: Plenum Press, pp. 387-402.

189. Moolgavkar, S.H. (1986) Hormones and multistage carcinogenesis. *Cancer Surveys* 5:635-648.

190. Morgan, K.; Patterson, D.; Gross, E. (1983) Formaldehyde and the Nasal Mucociliary Apparatus. *Formaldehyde: Toxicology, Epidemiology and Mechanisms*. New York: Marcel Dekker Inc., pp. 193-209.

191. Morris, S.C. (1990) *Cancer Risk Assessment*. New York: Marcel Dekker, Inc.

192. Muir, C.S.; Higginson, J. (1985) Chapter 2: Role of Epidemiology in Identifying Chemical Carcinogens. *Handbook of Carcinogen Testing*. New Jersey: Noyes Publications, pp. 28-57.

193. National Academy of Sciences (1980) *Principles of Toxicological Interactions Associated with Multiple Chemical Exposures*. Washington, D.C.: National Academy Press.

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

195. National Research Council (NRC) (1986) Dose-Route Extrapolations: Using Inhalation Toxicity Data to Set Drinking Water Limits. *Drinking Water and Health*, Vol 6. Washington, D.C.: National Academy Press.

196. National Research Council (NRC) (1987) *Pharmacokinetics in Risk Assessment: Drinking Water and Health, Volume 8.* Washington, D.C.: National Academy Press.

197. National Research Council (NRC) (1988) *Health Risks of Radon and Other Internally Deposited Alpha Emitters.* BEIR IV, Washington, D.C.: National Academy Press.

198. National Research Council (NRC) (1989) *Drinking Water and Health, Volume 9: Selected Issues in Risk Assessment.* Washington, D.C.: National Academy Press.

199. National Research Council (NRC) (1991) *Comparative Dosimetry of Radon in Mines and Homes.* Washington, D.C.: National Academy Press.

200. National Toxicology Program (NTP). (1984) *Report of the NTP ad hoc Panel on Chemical Carcinogenesis Testing and Evaluation.* U.S. Department of Health and Human Services, Washington, D.C.

201. Nesnow, S.; Bergman, H. (1988) An Analysis of the Gene-Tox Carcinogen Data Base. *Mutation Research* 205:237-253.

202. Newton-Smith, W. (1981) *The Rationality of Science.* London: Routledge and Kegan Paul.

203. OSHA (1980) Identification, Classification and Regulation of Potential Occupational Carcinogens. *Federal Register* 45:5129.

204. OSHA (1980) Occupational Exposure to Formaldehyde. *Federal Register* 50:50458.

205. Office of Science and Technology Policy (OSTP) (1985) Chemical Carcinogens: Review of the Science and its Associated Principles. *Federal Register* 50:10372-10442.

206. Park, C.N.; Kociba, R.J. (1985) Chapter 17. Statistical Evaluation of Long-Term Animal Bioassays for Carcinogenicity. *Handbook of Carcinogen Testing.* New Jersey: Noyes Publications, pp. 358-371.

207. Paustenbach, D.J.; Clewell III, H.J.; Gargas, M.L.; Andersen, M.E. (1988) A Physiologically Based Pharmacokinetic Model for Inhaled Carbon Tetrachloride. *Toxicology and Applied Pharmacology* 96:191-211.

208. Perera, F. (1987) The Potential Usefulness of Biological Markers in Risk Assessment. *Environmental Health Perspectives* 76:141-145.

209. Perera, F.P. (1984) The Genotoxic/Epigenetic Distinction: Relevance to Cancer Policy. *Environmental Research* 34:175-191.

210. Peto, J. (1984) Early- and Late-Stage Carcinogenesis in Mouse Skin and in Man. *Models, Mechanisms and Etiology of Tumour Promotion: IARC* 56:359.

211. Peto, R. (1985) Keynote Address: Epidemiological Reservations about Risk Assessment. *Basic Life Sciences, Vol. 33: Assessment of Risk from Low-Level Exposure to Radiation and Chemicals.*

212. Peto, R.; Pike, M.C.; Day, N.E.; Gray, R.G.; Lee, R.N.; Parish, S.; Peto, J.; Richards, S.; Wahrendork, J. (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 Man.*

213. Piegorsch, W.W.; Hoel, D.G. (1988) Exploring Relationships Between Mutagenic and Carcinogenic Potencies. *Mutation Research* 196:161-175.

214. Pitot, H.C.; Campbell, H.A. (1987) An Approach to the Determination of the Relative Potencies of Chemical Agents During the Stages of Initiation and Promotion in Multistage Hepatocarcinogenesis in the Rat. *Environmental Health Perspectives* 76:49-56.

215. Portier, C.; Hoel, D. (1983) Optimal Design of the Chronic Animal Bioassay. *Journal of Toxicology and Environmental Health* 12:1-19.

216. Portier, C.J. (1988) Species Correlation of Chemical Carcinogens. *Risk Analysis* 8(4):551-553.

217. Portier, C.J.; Bailer, A.J. (1989) Testing for Increased Carcinogenicity Using a Survival-Adjusted Quantal Response Test. *Fundamental and Applied Toxicology* 12:731-737.

218. Prejean, J.D. (1985) Chapter 13. Conduct of Long-Term Animal Bioassays. *Handbook of Carcinogen Testing.* New Jersey: Noyes Publications, pp. 268-281.

219. Purchase, I. (1985) Carcinogenic Risk Assessment: A Toxicologists View. *Risk Quantitation and Regulatory Policy: Banbury Report* 19:175-186.

220. Purchase, I.F.H.; Stafford, J.; Paddle, G.M. (1985) Chapter 8: Vinyl Chloride--A Cancer Case Study. *Toxicological Risk Assessment*, Vol. 2. CRC Press, pp. 167-194.

221. Ramsey, J.; Andersen, M. (1984) A Physiologically Based Description of the Inhalation Pharmacokinetics of Styrene in Rats and Humans. *Toxicology and Applied Pharmacology* 73:159.

222. Redmond, C.K. (1981) Sensitive Population Subsets in Relation to Effects of Low Doses. *Environmental Health Perspectives* 42:137-140.

223. Reitz, R.; Fox, T.; Quast, J. (1982) Mechanistic Considerations for Carcinogenic Risk Estimations: Chloroform. *Environmental Health Perspectives* 46:163.

224. Rose, M.; Lock, E.; Smith, L.; Wyatt, I. (1976) Paraquat Accumulation: Tissues and Species Specificity. *Biochemical Pharmacology* 25:419.

225. Rosin, M.; Stich, H. (1978) The Inhibitory Effect of Cysteine on the Mutagenic Activities of Several Carcinogens. *Mutation Research* 54:73.

226. Samet, J. (1989) Radon and Lung Cancer. *Journal of the National Cancer Institute* 81:745-755.

227. Saracci, R. (1977) Asbestos and Lung Cancer: An Analysis of Epidemiological Evidence on the Asbestos-smoking Interaction. *International Journal of Cancer* 20:323.

228. Sato, A.; Nakajima, T. (1979) Partition Coefficients of Some Aromatic Hydrocarbons and Ketones in Water, Blood and Oil. *British Journal of Industrial Medicine* 36:231.

229. Scala, J.; McOsker, D.; Reller, H. (1968) The Percutaneous Absorption of Ionic Surfactants. *Journal of Investigative Dermatology* 50:371.

230. Scherer, E.; Feringa, A.; Ennelot, P. (1984) Initiation-Promotion-Inititaion: Induction of Neoplastic Foci within Islands of Precancerous Liver Cells in Rats. *Models, Mechanisms and Etiology of Tumour Promotion: IARC 56.*

231. Schulte-Hermann, R.; Bursch, W.; Parzefall, W. (1991) Mitogenesis and Programmed Cell Death as Determinants of Carcinogenicity of Nongenotoxic Chemicals. *Chemically Induced Cell Proliferation: Implications for Risk Assessment.* New York: Wiley and Sons, pp. 237-244.

232. Schwartz, J.; Riddiourgh, C.; Epste, S. (1982) Analysis of Carcinogenesis Dose-Response Relations with Dichotomous Data. *Teratogenesis, Carcinogenesis and Mutagenesis* 2:179.

233. Schwetz, B.A.; Yang, R.S.H. (1990) Approaches used by the US National Toxicology Program in assessing the toxicity of chemical mixtures. *Complex Mixtures and Cancer Risk*, pp. 113-120.

234. Scribner, H.E.; McCarthy, K.L.; Doolittle, D.J. (1987) Practical Approaches to Evaluating Nongenotoxic Carcinogens. *Nongenotoxic Mechanisms in Carcinogenesis: Banbury Report* 25:355-364.

235. Shelby, M.D.; Stasiewicz, S. (1984) Chemicals Showing No Evidence of Carcinogenicity in Long-Term, Two-Species Rodent Studies: The Need for Short-Term Test Data. *Evnironmental Mutagen*, in press.

236. Short, B.; Swenberg, J. (1990) Chemically Induced Cell Proliferation in Carcinogenesis in the Male Rat Kidney. *Chemically Induced Cell Proliferation: Implications for Risk Assessment.* New York: Wiley and Sons, pp. 357-368.

237. Shubik, P.; Clayson, D.B. (1990) Chapter 8: Carcinogenesis in Animals: Which Agents Pose the Greatest Problems for Humans? *Progress in Predictive Toxicology*. New York: Elsevier Science Publishers, pp.145-160.

238. Sjoqvist, F.; Borga, O.; Orme, M. (1976) Fundamentals of Clinical Pharmacology. *Drug Treatment: Principles and Practice of Clinical Pharmacology and Therapeutics*. Australia: AIDS.

239. Slaga, T.J. (1984) Can Tumour Promotion be Effectively Inhibited? *Models, Mechanisms and Etiology of Tumour Promotion: IARC* 56:497-506.

240. Slaga, T.J. (1988) Interspecies Comparisons of Tissue DNA Damage, Repair, Fixation, and Replication. *Environmental Health Perspectives* 77:73-82.

241. Smith, R. (1973) *The Excretory Function of Bile: The Elimination of Drugs and Toxic Substances in Bile*. New York: Wiley & Sons.

242. Sorsa, M.; Vainio, H.; McMichael, A.J. (1990) Experimental and Epidemiologic Applications to Cancer Risk Assessment of Complex Mixtures. *Genetic Toxicology of Complex Mixtures*. New York: Plenum Press.

243. Spitzer, W.O.; Lawrence, V.; Dales, R.; Hill, G.; Archer, M.C.; Clark, P.; Abehnaim, L.; Hardy, J.; Sampalis, J.; Pinfold, S.P.; Morgan, P.P. (1990) Links Between Passive Smoking and Disease: A Best-Evidence Synthesis. *Clinical and Investigative Medicine* 13(1):17-42.

244. Standaert, F.G. (1988) Absorption and Distribution of Xenobiotics. *Environmental Health Perspectives* 76:63-71.

245. Starr, T.; Buck, R. (1984) The Importance of Delivered Dose in Estimating Low-dose Cancer Risk from Inhalation Exposure to Formaldehyde. *Fundamental and Applied Toxicology* 4:740.

246. Starr, T.; Gibson, J.; Swenberg, J. (1985) An Integrated Approach to the Study of Formaldehyde Carcinogenicity in Rats and Mice. *Toxicological Risk Assessment, Vol. II.* Florida: CRC Press, pp. 195-209.

247. Starr, T.B. (1990) Quantitative Cancer Risk Estimation for Formaldehyde. *Risk Analysis* 10(1):85-91.

248. Stowers, S.J.; Maronpot, R.R.; Reynolds, S.H.; Anderson, M.W. (1987) The Role of Oncogenes in Chemical Carcinogenesis. *Environmental Health Perspectives* 75:81-86.

249. Strom, S.C. (1987) Issues in Biochemical Applications to Risk Assessment: Can *In Vitro* Studies Assist Us in Species Extrapplation? *Environmental Health Perspectives* 76:181-184.

250. Swenberg, J.; Gross, E.; Martin, J., Popp, J. (1983) Mechanisms of Formaldehyde Toxicity. *Formaldehyde Toxicity*. Washington, D.C.: Hemisphere Publishers, pp. 132-137.

251. Swenberg, J.A.; Richardson, F.C.; Boucheron, J.A.; Deal, F.H.; Belinsky, S.A.; Charbonneau, M.; Short, B.G. (1987) High- to Low-Dose Extrapolation: Critical Determinants Involved in the Dose Response of Carcinogenic Substances. *Environmental Health Perspectives* 76:57-63.

252. Swenberg, J.A.; Short, B.G. (1987) Influence of Cytoxicity on the Induction of Tumors. *Nongenotoxic Mechanisms in Carcinogenesis: Banbury Report* 25:151-161.

253. Tamplin, A.; Cochran, T. (1974) A Report on the Inadequacy of Existing Radiation Protection Standards Related to Internal Exposure of Man to Insoluble Particles of Plutonium and Other Alpha Emitting Hot Particles, Natural Defense Council, Washington, D.C.

254. Tennant, R.W. (1987) Issues in Biochemical Applications to Risk Assessment: Are Short-Term Tests Predictive of *In Vivo* Tumorigenicity? *Environmental Health Perspectives* 76:163-167.

255. Tennant, R.W.; Ashby, J. (1991) Classification According to Chemical Structure, Mutagenicity to Salmonella and Level of Carcinogenicity of a Further 39 Chemicals Tested for Carcinogenicity by the US NTP. *Mutation Research* 257:209-227.

256. Tennant, R.W.; Margolin, B.H.; Shelby, M.D.; Zeiger, E.; Haseman, J.K.; Spalding, J.; Caspary, W.; Resnick, M.; Stasiewicz, S.; Anderson, B., Minor, R. (1987a) Prediction of Chemical Carcinogenicity in Rodents from in Vitro Genetic Toxicity Assays. *Science* 236:933-941.

257. Tennant, R.W.; Spalding, J.W.; Stasiewicz, S.; Caspary, W.D.; Mason, J.M.; Resnick, M.A. (1987b) Comparative Evaluation of Genetic Toxicity Patterns of Carcinogens and Noncarcinogens: Strategies for Predictive Use of Short-Term Assays. *Environmental Health Perspectives* 75:87-95.

258. Thorslund, T. (1990) Alternative Approaches for Estimating the Yearly Number of Lung Cancer Deaths in Non-Smokers due to ETS Based on Dose-Response Modeling. *Health Effects of Passive Smoking: Assessment of Lung Cancer in Adulst and Respiratory Disorders in Children.* US EPA, EPA/600/6-90/006A, Appendix D.

259. Toulmin, S. (1958) *The Uses of Argument Analysis.* Cambridge: Cambridge University Press.

260. Trosko, J.E.; Chang, C. (1988) Nongenotoxic Mechanisms in Carcinogenesis: Role of Inhibited Intercellular Communication. *Carcinogen Risk Assessment: New Directions in the Qualitative and Quantitative Aspects: Banbury Report* 31:139-159.

261. United States Department of Health and Human Services (1984) *Report of the NTP Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation.*

262. United States Environmental Protection Agency. (1983a) Good Laboratory Practices Standards--Toxicology Testing. *Federal Register* 48:53922.

263. United States Environmental Protection Agency. (1983b) Hazard Evaluations: Humans and Domestic Animals. Subdivision F. Available from: NTIS, Springfield, VA. PB 83-153916.

264. United States Environmental Protection Agency. (1983c) Health Effects Test Guidelines. Available from: NTIS, Springfield, VA. PB 83-232984.

265. United States Environmental Protection Agency (1984) Formaldehyde: Determination of Significant Risk. *Federal Register* 49:21874.

266. United States Environmental Protection Agency. (1986) Guidelines for Estimating Exposures. *Federal Register* 51(185):34042-34054.

267. United States Environmental Protection Agency. (1987) The Risk Assessment Guidelines of 1986. EPA/600/8-87/045.

268. United States Environmental Protection Agency. (1989) Biological Data for Pharmacokinetic Modeling and Risk Assessment. Report of a Workshop Convened by the U.S. EPA and ILSI Risk Science Institute. EPA/600/3-90/019.

269. United States Envrionmental Protection Agency (1990) Health Effects of Passive Smoking: Assessment of Lung Cancer in Adults and Respiratory Disorders in Children. USEPA/600/6-90/006A.

270. United States Food and Drug Admnistration (U.S. FDA) (1982) Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food. Available from: Bureau of Foods, U.S. FDA.

271. Vainio, H.; Hietanen, E. (1980) Role of Extrahepatic Metabolism in Drug Disposition and Toxicity. *Concepts in Drug Metabolism.* New York: Marcel Dekker Inc., pp. 251-284.

272. VanGinnekan, C.; Russel, F. (1989) Saturable Pharmacokinetics in the Renal Excretion of Drugs. *Clinical Pharmacokinetics* 16:38.

273. Vessell, E.; Penno, M. (1983) Intraindividual and Interindividual Variations. *Biological Basis of Detoxification.* New York: Academic Press.

274. Wahrendorf, J. (1990) Statistical Aspects of Experimental Studies with Complex Mixtures. *Complex Mixtures and Cancer Risk, IARC 1990.*

275. Wald, N.J.; Doll, R. (eds.) (1985) *Interpretation of Negative Epidemiological Evidence for Carcinogenicity.*

276. Weisburger, J. (1988) Cancer Risk Assessment Strategies Based on Mechanisms of Action. *Journal of the American College of Toxicology* 7:417-425.

277. Weiss, N. (1981) Inferring Causal Relationships: Elaboration of the Criterion of "Dose-Response". *American Journal of Epidemiology* 113:487-490.

278. Wheatley, D. (1968) Enhancement and Inhibition of the Induction by 7, 12-dimethylbenz(a)anthracene of Mammary Tumors in Female Sprague-Dawley Rats. *British Journal of Cancer* 22:787.

279. Whitey, J. (1982) Toxicodynamics and Biotransformation. *Assessment of Multichemical Contamination*. Washington, D.C.: National Academy Press.

280. Whitey, J. (1985) Chapter 2: Pharmacokinetic Differences Between Species. *Toxicological Risk Assessment Volume I: Biological and Statistical Criteria*. Florida: CRC Press, Inc., pp. 41-52.

281. Wilbourn, J.; Haroun, L.; Heseltine, E.; Kaldor, J.; Partensky, C.; Vainio, H. (1986) Response of Experimental Animals to Human Carcinogens: An Analysis Based upon the IARC Monographs Programme. *Carcinogenesis* 7(11):1853-1863.

282. Williams, G.M. (1984) Modulation of Chemical Carcinogenesis by Xenobiotics. *Fundamental and Applied Toxicology* 4:325-344.

283. Williams, G.M. (1987) The Significance of Preneoplastic Liver Lesions in Experimental Animals. *Advances in Veterinary Science and Comparative Medicine* 31:21-44.

284. Williams, G.M. (1990) Epigenetic Mechanisms of Liver Tumor Promotion. *Mouse Liver Carcinogenesis: Mechanisms and Species Comparisons*, pp. 131-145.

285. Williams, G.M.; Weisburger, J.H. (1991) Chapter 5: Chemical Carcinogenesis. *Toxicology*. New York: Pergamon Press, pp. 127-200.

286. Williams, R. (1974) Interspecies Variations in the Metabolism of Xenobiotics. *Biochemical Society Transactions* 2:359.

287. Withey, J.; Collins, B.; Collins, P. (1983) Effect of Vehicle on the Pharmacokinetics and Uptake of Four Halogenated Hydrocarbons form the Gastronintestinal Tract of the Rat. *Journal of Applied Toxicology* 3:249-253.

288. Wogan, G.; Gorelick, N. (1985) Chemical and Biochemical Dosimetry of Exposure to Genotoxic Chemicals. *Environmental Health Perspectives* 62:5.

289. Woodruff, M. (1990a) Chapter 1: Diversity of Cells in Tumours. *Cellular Variation and Adaptation in Cancer*. New York: Oxford University Press, pp. 1-7.

290. Woodruff, M. (1990b) Chapter 2: Transformation, Carcinogenesis, and Progression. *Cellular Variation and Adaptation in Cancer*. Oxford: University Press, pp. 8-27.

291. Woutersen, R.; Appleman, L.; Van Garderen-Hoetmer, A.; Feron, V. (1986) Inhalation Toxicity of Acetaldehyde in Rats. III. Carcinogenicity Study. *Toxicology* 41:213-231.

292. Zeiger, E.; Haseman, J.K.; Shelby, M.D.; Margolin, B.H.; Tennant, R.W. (1990) Evaluation of Four In Vitro Genetic Toxicity Tests for Predicting Rodent Carcinogenicity: Confirmation of Earlier Results With 41 Additional Chemicals. *Environmental and Molecular Mutagenesis* 16(18):1-14.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA/600/R-94/204a		

4. TITLE AND SUBTITLE	5. REPORT DATE
Hazard Identification in Carcinogen Risk Analysis: An Integrative Approach	October 1992
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Douglas J. Crawford-Brown, Kenneth G. Brown	OHEA-C-554 ; NCEA-W-013

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
Kenneth G. Brown, Ph.D. Inc. P.O. Box 16608 Chapel Hill, NC 27516-6608	11. CONTRACT/GRANT NO. 68-C9-0009; W.A. S-1-56

12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
National Center for Environmental Assessment* Washington Laboratory U.S. EPA Washington, DC 20460	14. SPONSORING AGENCY CODE EPA/600/021

15. SUPPLEMENTARY NOTES
This project report is part of the U.S. EPA's program, Research to Improve Health Risk Assessment (RIHRA). *Formerly Ofc. of Health & Environmental Assessment

16. ABSTRACT

The process of identifying carcinogenic agents, as the first stage in risk assessment, is examined from the point of view of rationality in a two-part report. The report is designed to aid the risk assessor in making rational judgments and to furnish a basis for discussion of these judgments with others, rather than attempting to formulate decision rules to avoid human decisions and avoid conflicts. The first part examines what is meant by the rationality of statements claiming that an agent is carcinogenic and enumerates several strategies by which one can claim that observations made in the laboratory or in human studies are relevant to human carcinogenicity. The second part of the report describes a seven-step framework that can be followed by an analyst in formulating judgments about hazard identification. This framework is based on the principles of rationality, coherence, and completeness discussed in the first part. Using this framework should accomplish three things: 1) aid in the systematic evaluation of all sources of information; 2) identify sources of divergence resulting from different perspectives and opinions of individuals; 3) help identify areas for research and help assess the potential impact of research on the hazard identification process.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
Release to the Public	Unclassified	225
	20. SECURITY CLASS *(This page)*	22. PRICE
	Unclassified	