J. Toxicol. Pub. Health
Vol. 17, Supplement, pp. 219-225, 2001

An Examination of Variation in Risk Assessment Practices in Relation to Assessors' Goals: American and International Practices

Lorenz R. Rhomberg*

Gradient Corporation, Cambridge, Massachusetts, USA

ABSTRACT *: The basic structure for assessment of potential health risks from environmental chemicals is widely agreed upon, but many of the details of risk assessment procedures differ among practitioners. Government regulatory agencies typically have guidelines or standard procedures for their risk assessments, established to ensure consistency and comparability, to set standards for adequacy, and to embody underlying tenets. In setting and updating such guidelines, each agency takes into account not only the prevailing thinking about appropriate procedures, but also its own goals and responsibilities and the precedents it has set for itself in past analyses. This results in variations in methods, and consequently in characterization of risks, among regulatory assessments, even when they are based on the same data. As a result, adopting existing assessments from a variety of regulatory bodies needs to be done with caution. This paper examines some of the variants in risk assessment approaches among American federal regulatory agencies and relates them to the variations in regulatory responsibilities of those groups. Comparisons to international practices are also drawn. The impact on development of world-wide risk standards is discussed.*

Key Words *: Risk Assessment Methodology, Harmonization, Non-Tariff Trade Barriers, Weight of Evidence, Carcinogenic Potency, Acceptable Daily Intake*

I. INTRODUCTION

In the last decades, the potential for exposures to pollutants and industrial chemicals to cause adverse impacts on human and ecological health has been of increasing concern throughout the world. An ever larger number of organizations carry out health risk assessments to aid in understanding potential environmental contamination problems, to set priorities for dealing with them, and to help decide on the actions, controls, and clean-ups that may be warranted. These bodies include governmental regulatory organizations, national and international scientific and public health organizations, industries, and others. In the United States, as in many nations, there are several government agencies at the national level that have differing but somewhat overlapping regulatory responsibilities, and there are many more such agencies at the state and local levels of government. Many of the responsibilities of such agencies are also addressed by supra-national organizations or are matters falling under international agreements on trade or environmental protection.

As a result of this multiplicity of organizations examining potential health and safety risks from environmental agents, the examination of particular agents may arise many times. A full risk analysis is expensive in terms of time, money, and the needed expertise and scientific infrastructure, and many organizations with limited resources may accordingly rely on assessments done by larger national and international organizations. Nonetheless, at best, a good deal of duplication of effort may result as each risk assessing organization goes through the process of adopting its own findings and judgments. At worst, there may be profound disagreements among assessments as to the nature and magnitude of health risks that an agent may pose, resulting from differences among assessments in methodology or scientific judgment. Especially in cases where regulatory responsibilities overlap, such variation among assessments can lead to conflict and uncertainty among regulators, the regulated community, and the public. This in turn can

*To whom correspondence should be addressed
Abbreviations : NAS, National Academy of Sciences; U.S., United States

undermine the confidence in the risk assessment process and in the legitimacy of environmental protection efforts.

In this paper, I examine some of the reasons for variations among assessments and discuss these in terms of the prospects for generating a base of universally recognized characterizations of risk. I base my discussion on findings of a report I wrote (Rhomberg, 1997) as a consultant to a special panel convened in the United States to examine risk assessment methodology as applied to environmental regulation: the Presidential/Congressional Commission on Risk Assessment and Risk Management. (The Commission's reports and my paper are available on the internet at <www.riskworld.com/Nreports/1996/risk_rpt/Rr6me001.htm>.) My purpose is not to tabulate the specific methodological differences among regulatory programs in the U.S. or worldwide. Such reviews are available in the literature (Carnevale *et al.*, 1987; Rosenthal *et al.*, 1992; Moolenar 1994; Schierow 1994; Sadowitz and Graham 1995; Hattis and Minkowitz 1995; Sanner *et al.* 1996; Rhomberg 1997; Beck *et al.* 2000; Seeley *et al.* 2000). Instead, I will discuss the general reasons for the existence of differences and the impact of their existence.

II. NEED FOR AGREEMENT AMONG ASSESSMENTS

Despite the difficulties of realizing widespread agreement among risk analyses, there are some needs that such agreement would address. These are best discussed by naming the difficulties that arise from lack of such agreement.

Without a dependably consistent basis for analysis, it is difficult to use the array of assessments that have been produced on different agents and environmental problems in assessing comparative risks, risk tradeoffs, and the setting of environmental protection priorities. If an organization wishes to compare risks across such an array, it needs to expend considerable efforts to ensure that all assessments adhere to a common set of assumptions and assessment methods. The constant revisiting of risk questions by each interested organization is wasteful and beyond the meager resources of all but the largest groups. The use of existing assessments by authorities with fewer resources to expend on generating their own is hampered. Inconsistencies in data requirements mean that companies attempting to comply with regulations in many different political jurisdictions must often undertake essentially duplicative testing to ensure compliance with each set of standards, also wasting resources and inhibiting international development of markets. Differing assessments can be used as the basis of non-tariff trade barriers, or the accusation that such use is being made can cause conflicts in international trade.

At the same time, it must be acknowledged that risk assessment represents coming to scientific judgments about the application of toxicological and epidemiologic data to the assessment of potential health risks in populations of concern. It is the right of different entities to come to their own judgments based on their own criteria for sufficiency of evidence, appropriate scientific interpretation, and the casting of their findings so as to be most useful to their particular risk management problems.

To better understand these conflicting principles, it is useful to consider the reasons for variation in risk assessment methodology and in outcomes of examination of particular agents. These flow from the nature of environmental health risk assessment.

III. THE NATURE OF RISK ASSESSMENT

Risk assessment can be thought of as bringing to bear existing scientific information—and the critical interpretation of that information—on questions of the existence, nature, and magnitude of risks that many be posed by exposure to an agent. The difficulties are that information is always incomplete and inconclusive to some degree. The scientific basis is often indirect (based on animal studies or on highly exposed occupational populations, for example) and one must extrapolate and generalize the available findings, often to a considerable degree, to apply the results to the environmental exposures of concern. Inherently unobservable phenomena (low-level risks that cannot be detected epidemiologically) are of concern. Often, available data are apparently contradictory, and alternative scientific explanations (with different risk consequences) are possible. All of this is conducted as a public process in the face of con-

flicting interests, and interested parties argue that risk management solutions are being unfairly imposed upon them based on uncertain science.

The essential problem is that the needs of the risk assessment process far outstrip the ability of scientific investigation to give firm answers, yet firm answers seem to be required by the regulatory process to take concrete actions. The practical need remains to make characterizations of the risk consequences of various actions and activities. Faced with this practical problem, regulatory agencies have developed risk assessment methods that, while attempting to embody available information, of necessity rely on uncertainty-bridging principles derived from a combination of general knowledge about chemicals and their toxic effects. The specification of such methods is aimed at ensuring that consistency is maintained case to case, that uncertainties are resolved by common logic, and that the decisions made fulfill the mandates for public health protection that the assessments must support. It is inconsistency in such rules—or in their application from case to case—that leads to variation in risk analyses even when based on the same set of toxicological and epidemiologic data.

Because application of judgment is necessary and data do not simply speak for themselves, there is concern that political and economic considerations can color the conclusions about toxicity. To defend against this problem, the U.S. National Academy of Sciences proposed a structure for conducting risk analysis (NAS, 1983) that called for risk assessing government agencies to produce guidelines and policy documents that outline and explain methodology and provide a consistent set of principles for the application of methods, rules of evidence, and necessary assumptions. The NAS sought to impose some structure on the risk assessment process, noting that it has qualitative aspects—the determination of the specific kinds of toxic effects that an agent may be capable of causing, and an assessment of the certainty that effects seen in animals apply to humans—and quantitative aspects—the estimation of how the probability or magnitude of response will vary according to the amount of agent experienced. The NAS further suggested that one should distinguish the characterization of potential risks from the questions about what should be done to control or mitigate them. That is, one aims at an objective characterization of what might be caused by exposure to an agent, and only then applies the considerations of costs, benefits, social impacts, and so on that are a legitimate part of societal decision-making about how to handle risky activities.

Although this exact framework is not universally applied, the distinctions it makes are widely recognized, and the basic structure for risk assessment of potential health risks from environmental chemicals is widely agreed upon. Nonetheless, many details differ among practitioners. The implementation of explicit guidelines has been uneven; in some bodies, comprehensive methodology is laid out but in others the methods rely on the body of precedent and conventional practice.

IV. REGULATORY USES OF RISK ASSESSMENT

Risk analyses can be applied to a variety of problems, including permitting and licensing, supporting regulation of emissions, disposal, and handling, assessing safety of existing practices, analyzing existing environmental problems and setting remediation requirements, and priority setting. The mandates and responsibilities that a regulatory agency may be given in its authorizing laws are typically phrased in very general terms. Considering examples from the U.S., the Federal Food Drug and Cosmetic Act calls for "reasonable probability of no harm"; the Clear Air Act requires that the Environmental Protection Agency "protect the public health with an adequate margin of safety"; the Consumer Product Safety Act has a mandate to "protect the public against unreasonable risks"; and the Occupational Safety and Health Act demands actions to "assure insofar as practicable that no employee will suffer diminished health, functional capacity, or life expectancy". Although these phrases are general, they differ subtly in their requirements for the degree of assurance of safety required, the toleration of uncertainty, and the allowance for considerations of costs and practical difficulties of control. Some are addressed to control of future activities, some to remediation of past contamination. Some are aimed at exposures experienced by

the national population, others at exposures to relatively few people in specific locations or occupations. Some address exposures encountered in a voluntary way, others deal with exposures that the populace may be exposed to against its will. Although the laws do not (usually) explicitly call for specific alterations in risk assessment methodology, it is easy to see that these different purposes can legitimately have different tolerances for uncertainty, needs for robustness of supporting data, and different embedded assumptions about the nature and magnitude of typical exposures.

It is up to each regulatory agency to achieve its mandate by employing the regulatory powers it is given. These also differ considerably from setting to setting. The regulatory tools available may be the ability to grant or deny licenses, permit or ban use, control distribution or use of an agent, specify conditions for use or protective equipment to be employed, classify agents into categories of permitted activities, or labeling and provision of warnings, but each regulatory setting allows only a few of these, and the conduct of the analysis will also be influenced by what the agency can do to achieve its mandate. This is not to say that the results of the risk analysis will be modified to suit regulatory preferences, but rather that the nature of the risk management questions changes among applications, and the most informative risk assessment practices will also change accordingly. For instance, when a regulatory threshold exists (*e.g.*, label or do not label), the assessment may be geared to provide enough detail only to settle the immediate issue, not to estimate the residual risks that may occur depending on different actions.

The risk assessment process requires considerable scientific judgment, not only in interpretation of data bearing on an agent's potential risks and its extrapolation to the situations of concern, but also in the definition of appropriate default assumptions and extrapolation methods. When these judgments are made by different bodies, it is natural that some legitimate differences of opinion will manifest themselves, over and above those just mentioned in connection with different regulatory purposes and powers. To some degree, risk assessment methods differ owing to different histories and precedents about how such judgments have been made. Once practices have been set up, they tend to be followed in future cases unless there is compelling reason not to do so. This is for purposes of consistency and comparability. Such consistency is laudable, but when patterns of practice originate in several independent organizations, the chosen paths may differ from one another for nothing other than historical reasons. Once these paths are set, changes seem disruptive and harmful to the goals of consistency within programs.

V. REASONS FOR VARIATION IN METHODS

I suggest that the reasons for variation among regulatory agencies in risk assessment methodology can be broadly grouped under the following headings: Purpose, Basis, Rules, Understanding, Precaution, and Process.

Purpose - As noted, different organizations have different regulatory mandates. Their stated aims may be to ensure against risk with great certainty, to provide margins of safety, to balance costs and benefits, to define reasonable precautions, to warn consumers, to protect populations or individuals. The exposed populations differ in size, nature, and the degree to which protection from impact can be expected. The tools they can use may be bans, controls, restrictions, or labels. The risk assessment approaches that best address these different needs will be somewhat different in their tolerance of uncertainty, need for rigor in data and extrapolations, need to make central estimates or bounds, focus on individual risks or population risks, consideration of variation in sensitivity, assumptions about concomitant exposure, and so on, and methods chosen may vary accordingly. When an assessment is brought to bear on a question other than the one it was originally formulated to address, the influence of such differences should be attended to.

Basis - There may be different rules for using toxicologic, epidemiologic, and mechanistic data among organizations, so it is not always so that different groups consider the same array of studies. Some organizations require that source material be published or peer reviewed, some require minimum standards of study design or monitoring of conduct, some assessments consider malignant tumors only while others consider malignant plus benign tumors as indicative of impacts, some require human data, oth-

ers give it preference, and still others treat human and animal data on equal footing, and so on. Depending on these rules of evidence, some studies may be ruled in by one assessment and out by another, leading to different interpretations of the array of findings on which to base inferences.

Rules - To achieve consistency, the standards for interpretation of information are frequently codified into rules, and even essentially similar judgments can appear different as a result of how categories are defined and how defaults are applied. For example, in hazard identification, the weight of evidence that an agent is a human carcinogen is often divided into distinct named compartments (known, likely, *etc.*), despite the fact that the judgment is really based on a continuum. Different organizations can place the cut-points for categories differently and thus express similar judgments in different terms. Some organizations focus on so-called strength-of-evidence (considering primarily positive findings) and others on weight-of-evidence (giving weight to both positive and negative indications of hazard). Certain data on presumed mode of toxic action can place an agent in one or another category that receives different methodological treatment (*e.g.*, genotoxic *vs.* non-genotoxic carcinogens). Some assessments consider potency or shape of the dose-response curve or margin of exposure in classifications, while others do not.

In quantitative risk assessment, the rules deal with default choices of dose-response model, fitting methods and criteria for acceptance or rejection of fitted equations, treatment of intercurrent mortality in lifetime bioassay data, means of expressing risk above background, and other details that can alter the quantitative description of particular toxicological testing results, quite aside from the questions about how such results are to be extrapolated to apply to low environmental exposures in humans. Rules for such extrapolations also differ, embodying somewhat different rationales for extrapolation and what is best presumed in the absence of case-specific data. The diversity of specific methods correctly reflects the uncertainty about what constitutes the best rationale for such extrapolations, and it would be hard to obtain agreement as to which particular approaches are best and unbiased. Such rules are particularly prone to the freezing by adherence to precedent previously mentioned. The result is that different organizations, even when starting with the same bioassay data, can end up with a host of minor differences and sometimes some major differences in presumed impact.

Understanding - To some degree, the diversity of rules discussed above reflects differences in understanding or interpretation of the general principles involved in arriving at defaults. There can also be differences in judgment about the bearing of case-specific data. For example, relevance to humans of certain modes of carcinogenic action seen in animal studies has been subject to debate, and different assessments can vary according to how they read the relevance of a particular bioassay result, *i.e.*, whether the relevance has been established or refuted with an appropriate degree of rigor.

Precaution - Uncertainty in risk analysis is inescapable, and methods for conducting assessments vary in how they express such uncertainty as well as the degree to which they allow for it. Approaches include the use of explicit uncertainty factors, the use of conservative values of parameters and bounds on extrapolations, the establishment of thresholds of evidence for departing from defaults, and the general use of the so-called precautionary principle. Most regulatory assessments include various of these in different combinations. In extrapolation, it is often rather unclear how much of the adjustments that are made are applied strictly for extrapolation (*i.e.*, to generalize a central estimate to another situation) and how much to allow for uncertainty in the extrapolation. This makes it difficult to "remove" allowances for uncertainty after the fact to render assessments done by different organizations more comparable.

Process - Aside from the content of a risk assessment, the steps in the bureaucratic process for establishing findings by a regulatory body vary. In some cases, national panels of experts make judgments that are essentially accepted without recourse. In others, government offices or ministries play such a role. There may be differing requirements for justifying actions by reference to a factual record or to explicit rules of procedure. In some settings, assessments are essentially negotiated by a group of stakeholders in the decision. There may be differing opportunities for public comment, appeal of decisions, and judicial

review of assessments and regulatory actions. In the United States, the role of the factual record, rules of evidence, adherence to stated guidelines that describe objective criteria, public comment and review, and recourse to judicial review of regulatory actions play a large role. In other settings, assessments are recognized more as a matter of judgment by competent experts. The result of these differences is that assessments conducted in different settings have different degrees of necessity to address concerns of a variety of parties, they differ in the degree to which decisions need to be defended in adversarial settings, and in the degree to which they need to be shown as flowing from defined criteria and methodology.

VI. CONCLUSIONS

For all of the reasons discussed above, risk assessments conducted by different organizations can vary, even when ostensibly based on the same set of toxicological, epidemiological, and mechanistic data. Some of the reasons why this is so can be decried as due to bureaucratic inertia, stubborn adherence to traditional approaches in the face of improving science, resistance to harmonization, and unwillingness to address specific cases on the case-specific evidence. But there is also a set of legitimate reasons why assessments vary, having to do with the purposes to which they are to be put and the risk management questions they are required to address. In any case, the notion that a single common "objective" assessment can be defined that only deals with facts and central estimates, eschewing assumptions, defaults, and conservative allowances for uncertainty, is not realistic. Every assessment entails a large degree of generalization from the specific facts arising in the context of specific studies to make conclusions about different settings of interest. Such extrapolations necessarily entail interpretations of the bearing of general understanding of biology and toxicology and the choice of principles on which to achieve the projection of interpretations that is necessary. The balance of general principles and case-specific data, and the standards of evidence needed to adopt special procedures that may be either astute use of compound-specific data or unproven speculations depending on one's viewpoint, is always an issue, and standards legitimately vary among circumstances. Forcing standardization of assessment approaches is unlikely to completely succeed, and in any case risks settling for the least common denominator and stifling innovation and creative case-specific approaches.

The best practice in considering assessments from different organizations is to interpret them in the context under which they were formulated. This is aided by the use of explicit guidance and articulation of assumptions, methods, and the justifications for them, and the production of such documentation is to be encouraged. Similarly, specific assessments should be explicit about the use of assumptions and variant methodologies, articulating the basis for their approaches so that the context can indeed be appreciated. The differences in method and meaning of assessments will continue to pose a challenge to the creation of a world-wide base of information on risks and potencies of toxic agents.

REFERENCES

Beck, B.D., Seeley, M.R., Tonner-Navarro, L.E. and Deskin, R. (2000): Variation in cancer classification between European countries and organizations. *The Toxicologist*, **54**, 1283.

Carnevale, F., Montesano, R., Partensky, C. and Tomatis, L. (1987): Comparison of regulations on occupational carcinogens in several industrialized countries. *Am. J. Ind. Med.*, **12**, 453-473.

Hattis, D. and Minkowitz, W.S. (1995): Risk evaluation: legal requirements, conceptual foundations, and practical experience in the United States, in Risk Evaluation–A Survey of International Experience (Renn, O. ed.) Kluwer Academic Publishers, Amsterdam.

Moolenar, R.J. (1994): Carcinogen risk assessment: international comparison. *Regul. Toxicol. Pharmacol.*, **20**, 302-336.

Rhomberg, L.R. (1997): A survey of methods for chemical health risk assessment among federal agencies. *Human Ecol. Risk Assessment*, **3**, 1029-1196.

Rosenthal, A., Gray, G.M. and Graham, J.D. (1992): Legislating acceptable cancer risk from exposure to toxic chemicals. *Ecol. Law Quart.*, **19**, 269-362.

Sadowitz, M. and Graham, J.D. (1995): A survey of residual cancer risks permitted by health, safety, and environmental policy. *Risk, Health, Safety & Environment*, **17**, 17-35.

Sanner, T., Dybing, E., Kroese, D. Roelfzema, H. and Hardeng, S. (1996): Carcinogen classification systems:

similarities and differences. *Regul. Toxicol. Pharmacol.*, **23**, 128-138.

Schierow, L.-J. (1994): Risk Analysis and Cost-Benefit Analysis of Environmental Regulations, CRS Report to Congress No. 94-961 ENR, Congressional Research Service, Library of Congress, Washington DC.

Seeley, M.R., Beck, B.D., Tonner-Navarro, L.E. and Deskin, R. (2000): European differences in derivation of occupational exposure levels. *The Toxicologist*, **54**, 1284.