Dr. Nancy Beck
Office of Information and Regulatory Affairs
Office of Management and Budget
725 17th Street, NW.
New Executive Office Building
Room 10201
Washington, DC 20503

June 15, 2006

Re: Comments on Proposed Risk Assessment Bulletin

Dear Dr. Beck:

The American Chemistry Council (ACC or the Council) is pleased to submit comments on the Office of Management and Budget's Proposed Risk Assessment Bulletin[1]. The Council represents the leading companies engaged in the business of chemistry[2].

ACC and its members make substantial, ongoing investments in research to support product development, health, safety and environmental protection, and to abide by product stewardship and regulatory policies. Chemistry is a science-based industry, and ACC has long sought to improve the quality of government science generally and risk assessment in particular. For example, in response to OMB's Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations, ACC filed an extensive set of comments (63 pages plus five appendices) that primarily focused on EPA's risk assessment practices.[3] Appendix 5 to those comments provided 62 additional pages of examples of EPA risk assessments that overstated risks. ACC's comments were the principal drivers behind EPA's 2004 staff paper on the Agency's risk assessment principles and practices – a document which defended the appropriateness of many of the practices to which ACC objected.[4] These controversies are still largely unresolved, and thus ACC has a substantial interest in the Bulletin.

---

[1] Notice of availability at 71 Fed. Reg. 2600 (Jan. 17, 2006).

[2] Council members apply the science of chemistry to make innovative products and services that make people's lives better, healthier and safer. The Council is committed to improved environmental, health and safety performance through Responsible Care®, common sense advocacy designed to address major public policy issues, and health and environmental research and product testing. The business of chemistry is a $460 billion enterprise and a key element of the nation's economy. It is the nation's largest exporter, accounting for ten cents out of every dollar in U.S. exports. Chemistry companies invest more in research and development than any other business sector.

[3] ACC, "Comments to the Office of Management & Budget; Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations," filed May 5, 2003. These comments and Appendix 5 are attached.

[4] EPA Office of the Science Advisor Staff Paper, *An Examination of EPA Risk Assessment Principles and Practices* (EPA/100/B-04/001) (Feb. 2004).

ACC has strongly supported OMB's efforts – through its Information Quality Act (IQA) Guidelines, [5] the Peer Review Bulletin, [6] Circular A-4,[7] and otherwise – to assure that the highest quality scientific work products are consistently and assiduously applied in support of regulatory policy. The proposed Risk Assessment Bulletin continues those efforts, and ACC applauds OMB for issuing it. We believe the Bulletin, once finalized, will improve the uneven performance of risk assessments at EPA and other federal agencies by setting a unified, upgraded standard.

The attached comments highlight the strengths that we have identified in the document, and recommend a number of improvements that we believe are vital to its success. We understand that many important issues associated with the Bulletin will only become clear as it is implemented, and we look forward to a continuing dialogue with OMB before and after its final publication. Should you or other OMB staff have any questions on, or need clarification of, ACC comments, please don't hesitate to contact either of us at 703-741-5000.

Sincerely,

James W. Conrad, Jr.
Assistant General Counsel

Richard A. Becker, Ph.D. DABT
Senior Toxicologist/Senior Director

Attachment: Comments of the American Chemistry Council on the Proposed Risk Assessment Bulletin (released for public review and comment in January, 2006)

[5] OMB, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*, 67 Fed. Reg. 8452 (Feb. 22, 2002).
[6] OMB, *Final Information Quality Bulletin for Peer Review*, 70 Fed. Reg. 2664 (Jan. 14, 2005).
[7] OMB, *Circular A-4* (Sept. 2003).

COMMENTS TO THE OFFICE OF MANAGEMENT AND BUDGET

DRAFT 2003 REPORT TO CONGRESS ON THE COSTS AND BENEFITS OF FEDERAL REGULATIONS

68 FR 5492

February 3, 2003

Submitted on

May 5, 2003

The American Chemistry Council

Arlington, Virginia

Table of Contents

I. INTRODUCTION

The American Chemistry Council ("Council") is pleased to comment on the Office of Management and Budget's ("OMB") Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations (68 FR 5492-5527), which was released for public comment on February 3, 2003.

The Council represents the leading companies engaged in the business of chemistry. Council members apply the science of chemistry to make innovative products and services that make people's lives better, healthier and safer. The Council is committed to improved environmental, health and safety performance through Responsible Care, common sense advocacy designed to address major public policy issues, and health and environmental research and product testing. The business of chemistry is a $460 billion enterprise and a key element of the nation's economy. It is the nation's largest exporter, accounting for ten cents out of every dollar in U.S. exports. Chemistry companies invest more in research and development than any other business sector.

In the year 2000, companies in the business of chemistry spent an estimated $19.7 billion to comply with all federal regulations. For these companies, this figure is equivalent to $19,000 per worker per year. The largest share (58%) is due to environmental regulation, followed by economic (15%), tax (14%), and workplace regulation (12%). Compliance with environmental regulation cost the business of chemistry $11.5 billion in the year 2000. Included in this category are regulations to control air and water pollution, reduce the risks posed by chemical products, and manage hazardous waste. About half of all environmental spending represents a recurring cost associated with pollution abatement control. About 30% of environmental spending is due to one-time capital costs. Hazardous waste cleanup represents 20% of environmental spending.

The Council recognizes that cost cannot serve as the sole basis for judging the value of federal regulations. Nevertheless, the magnitude of these costs raises serious questions about the cost-effectiveness of our nation's regulatory expenditures. For that reason, the Council welcomes OMB's requests for comments on the Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations ("Draft Report") and the Draft Guidelines for the Conduct or Regulatory

Analysis and the Format of Accounting Statements ("Draft Guidelines").[1] The Council particularly appreciates OMB's interest in receiving comments on the U.S. government's approaches to analysis and management of emerging risks and how the U.S. government balances precautionary approaches to health, safety and environmental risks with other interests such as economic growth and technological innovation. When risks are exaggerated rather than estimated scientifically and objectively, risk managers can be diverted away from managing meaningful risks, resulting in an inefficient use of resources that could be better directed toward higher priority problems. As our comments will indicate, we are concerned that current practices, especially at EPA, have followed this path.

The Council hopes these comments will assist OMB and others in the Executive Office seeking to improve federal regulatory analysis and management. The Council looks forward to working with OMB on this and similar matters in the future.

[1] The Draft Guidelines are appended to the Draft Report to Congress as Appendix C.

II. EXECUTIVE SUMMARY

As a fundamental principle, the Council and its members support health, safety, and environmental protection policies that incorporate risk-based priorities and cost-effective risk management. Essential to realizing such policies are well-conducted assessments of the costs and benefits of intended regulations, consistent with the regulatory principles of E.O. 12866, on "Regulatory Planning and Review." Moreover, as articulated in John D. Graham's September 20, 2001, memorandum on Presidential Review of Agency Rulemaking by OIRA, the risk assessments that are integral to this process should provide "an objective, realistic, and scientifically balanced analysis." Congress, in the Safe Drinking Water Act of 1996, also underscored as a national policy that agencies' science-based regulatory decisions for drinking water must use the "best-available, peer-reviewed science" and that agencies must present "comprehensive, informative, and understandable" information about the risks they regulate. OMB has recommended that agencies adopt or adapt these standards more broadly for judging the quality of scientific information they disseminate about risks they assess and regulate. Indeed, OMB has set high standards for scientific risk assessment for more than 15 years. Yet our experience is that EPA continues to issue guidelines and individual assessments that fall significantly short of the standards for objective, science-based, and realistic risk assessments (e.g., 2003 Draft Carcinogen Risk Assessment Guidelines and Supplement, EPA's chloroform risk assessment) through an over-reliance on highly conservative, "worst case" approaches, even for comprehensive assessments when more realistic information is available.

Even the most complete and data-derived risk analysis will require what the National Research Council's 1983 report, *Risk Assessment in the Federal Government: Managing the Process*, called "inferential bridges" (also, "risk assessment policy," etc., p3) or default assumptions to fill data gaps and scientific uncertainties. This does not mean, however, that it is acceptable to ultimately manage risk based on unjustified assumptions and policies that generate unrealistically biased and exaggerated risk assessments. Analyses that are purposely and sometimes highly conservative may be acceptable for early-tier screening assessments aimed at determining whether further investigation is needed, but they are not appropriate for end-stage

risk analysis that will drive risk management decisions. Yet – in a number of cases affecting Council members and other regulated entities – EPA later-stage risk assessments have relied on a cascade of conservative policy assumptions, despite more than a decade of attempts by OMB, Congress, and others to improve the Agency's practices.

The Council understands and supports the fundamental intention of OMB and the nation's regulatory system to achieve risk-based, cost-effective decisions, with appropriately applied precaution to deal with legitimate scientific uncertainties and data gaps. But, as more fully discussed in our comments – and detailed in Appendix 5 – the Council has deep concerns that decisions made in the name of environmental protection, and their supporting risk analyses, far too often embody an overly precautious, and often invisible, bias. Our comments provide not only the numerous specific examples cited in Appendix 5, but also provide a table in Appendix 4 listing the numerous conservative default assumptions embodied in EPA guidance and methodologies. We greatly appreciate OMB's initiative to better understand and remedy the continuing deficiencies in what should be an effective regulatory system that promotes the American public's interest in genuine health, safety, and environmental protection together with the economic prosperity and innovation that are the foundation for our other important goals.

The Council's comments respond to several parts of OMB's requests for comments:

- OMB's Draft Guidelines for the Conduct of Regulatory Analysis and the Format of Accounting Statements ("Draft Guidelines"), Appendix C.

- U.S. approaches to analysis and management of emerging risks, with a particular focus on ways in which "precaution" is embedded in current risk assessment procedures and examples of unbalanced approaches to human and ecological risk assessment.

- Briefly, we comment on the question of how the U.S. balances precautionary approaches to health, safety, and environmental risks with other interests.

While we do not provide any suggestions on how to conduct better analyses of regulations related to homeland security, we do agree that these regulations can have significant costs and benefits that should be evaluated with the same degree of care appropriate for other kinds of

regulations. In addition, we would refer you to the Council's comments of May 24, 2002, in which we discuss the need for agencies to standardize their approaches to valuing effects and also comment on other issues pertaining to OMB's Annual Report to Congress on the Costs and Benefits of Federal Regulations, but we do not discuss the Draft 2003 Report in these comments. A summary of our recommendations on the above three issues follows:

I. The Council finds OMB's Draft Guidelines for the Conduct of Regulatory Analysis to be well written, comprehensive, and strongly grounded in economic theory and practice. We support OMB's proposal to have agencies conduct both Benefit-Cost Analyses and Cost-Effectiveness Analyses of regulations as a way to gain important additional insight into the merits of regulatory proposals. We also agree that analysis of 3% and 7% discount rates in evaluating future benefits and costs is an appropriate starting point, but also believe that as a sensitivity analysis higher discount rates should also be used along with these rates. Indeed, as OMB notes, at times much higher rates would be plausible. The Council agrees and recommends OMB establish a 7% real discount rate as a "weak" default intended to permit ready comparison, but require that agencies perform sensitivity analyses across a very wide swath of discount rates that could be reasonable under particular circumstances for specific subpopulations.

II. In seeking public comment on current U.S. approaches to the analysis and management of emerging risks, OMB specifically asked about ways in which "precaution" is embedded in current risk assessment procedures and sought specific examples of unbalanced approaches to human and ecological risk assessment. While other federal regulations besides those issued by the EPA impact companies in the business of chemistry, the Council's comments focus almost exclusively on EPA risk regulations because these have the greatest impact on our members and represent the most clearly illustrated, indeed sometimes egregious, examples of regulations that could be made more realistic and cost-effective. Besides the specific examples provided in the appendix, the Council makes a set of

core policy recommendations aimed at improving the conduct of risk assessment and management, none of them original but all of them important and needing more rigorous implementation to be effective.

III. The Council emphasizes the need to consider risk management in the context of other public values, including a prosperous economy and innovation, which can be harmed by lopsided and extreme precaution.

The Council's core policy recommendations for improving EPA risk assessments involve three fundamental changes to the Agency's practice and several improvements in risk assessment procedures that will assist in realizing these fundamental changes. While it is vital for EPA to take the immediate steps described below to improve Agency risk assessments, the Council also urges OMB to work with EPA to set in motion a long-term process of revising its risk assessment guidelines and methodologies to make them current with the scientific state-of-the-art and the high standard OMB has articulated for risk analyses. As noted above, none of the Council's suggested changes are new – rather, they are based on basic tenets of a science-based risk assessment process that were formulated years ago[2] but that EPA has yet to adequately implement. The Executive Office of the President – in a 1991 document titled "Regulatory Program of the United States Government" (EOP, 1991) – set forth the fundamental benchmarks that EPA's risk assessment process must live up to. A panel of 15 invited experts and 35 other participants also reviewed this document and concluded, among other things, that the continued reliance on worst-case assumptions distorts risk assessments. [Evans 1992] Quotes and key ideas in the following three points are taken from the EOP 1991 document:

- *EPA risk assessments must not "intermingle important policy judgments within the scientific assessment of risk" Rather, the "choice of an appropriate margin of safety should remain the province of responsible risk-management officials, and should not be preempted through biased risk assessments.*" This principle is simple – risk assessments should aspire to the greatest extent possible to be objective scientific exercises that seek to realistically estimate risk. Risk management comes later, and

[2] See, e.g., NRC (1983); EOP (1991).

must be fully and transparently distinguished from risk assessment if the practice of risk assessment is to have scientific credibility.

- *Risk assessments should not continue an unwarranted reliance on "conservative (worst-case) assumptions" that distort the outcomes of the risk assessment, "yielding estimates that may overstate likely risks by several orders of magnitude*." Further, worst-case assumptions concerning actual human exposure should not be "used instead of empirical data," because they further exaggerate predicted risk levels. In short, risk assessments should use real data to the extent feasible; be as accurate, precise and realistic as possible; and should not seek to embed conservative policy preferences into what should be a policy-neutral estimation of risk. In cases where realistic knowledge/data concerning a risk scenario is unavailable, assumptions are often necessary, but to the extent that conservative assumptions are used in a risk assessment they must be clearly articulated for risk managers so that they fully understand how the analysis was performed, and where it may be overestimating risks.

- *Risk assessments should "acknowledge the presence of considerable uncertainty" and present the extent to which conservative assumptions may overstate likely risks*. They should not "routinely ignore these uncertainties and treat the resulting upper-bound estimates as reliable guides to the likely consequences of regulatory action." Risk assessments should directly assess the impact of each choice or assumption and clearly communicate how these choices impact likely risks.

As discussed below, the Council urges OMB to work with EPA to immediately strengthen several risk assessment procedures and tools to foster clearer adherence to the three fundamental principles. These are:

⇒ *Risk assessors should present managers with a range of risk scenarios and fully disclose the plausibility of each to facilitate the risk manager's informed policy choices*. OMB must direct agencies in their risk assessments to consider multiple scenarios and to fully account for the plausibility or likelihood of each. Within this process, agencies must consider the highly

unlikely but plausible worst-case, the expected value or mean estimate of risk, and the reasonable best-case outcomes, without unduly emphasizing worst-case hypothetical scenarios. In presenting risk managers with a range of possible outcomes, accurately weighted for their likelihood, the goal of risk assessors should be to support the managers in making fully informed choices about both the appropriate degree of conservatism or precaution to adopt and the extent to which such choices may entail tradeoffs among other important factors (i.e., to facilitate the risk manager's informed consideration of benefits and costs). In comprehensively disclosing the features of their assessments, risk assessors must provide the empirical basis or scientific rationale for any assumption, conservative or precautionary policy choices used in a given scenario. They must also fully explain the implications of choosing a particular policy, including the countervailing risks and other effects that might arise directly or indirectly from a decision based on such policy choices. While default assumptions are required to fill data gaps and address uncertainties that arise in the conduct of a risk assessment, it is the risk manager's responsibility to ultimately decide how to address limitations in the risk assessment through additional safety factors and other policy decisions. Risk assessments must serve, not usurp, this process. As discussed in greater detail under Section 6 of these comments, one specific tool that would support risk managers in their role is greater reliance on Monte Carlo and other stochastic methods in conducting risk assessments. EPA has endorsed such methods for exposure assessments [EPA 1997], but has not facilitated their use by defining the process and data to be used.

⇒ *Agencies should assess scientific evidence using a weight-of-the-evidence process that is consistent, comprehensive, balanced, and reproducible*. Although EPA often describes the process it uses in toxicity assessments (and sometimes in performing exposure assessments) as a weight-of-the-evidence approach, in fact the Agency does not follow consistent, comprehensive, balanced, and reproducible procedures that external parties, such as the Council, can follow and understand. Such procedures assist the risk assessor

in deciding which data, both positive and negative, should be given more weight, and in determining how disparate data can be combined to reach a rational and scientifically supportable conclusion. To be useful and understandable to external parties, EPA's assessments must employ a more formal and transparent weight-of-the-evidence process (for example, the approach developed by Klimisch, et al., for evaluating data quality [Klimisch, 1997], the Bradford-Hill causation criteria cited below, and other such approaches that can make it clearer how EPA risk assessors judged the evidence they considered). A formal process would assign weights to data or apply carefully defined evaluation criteria to assist the risk assessor in deciding what data should be given more weight and in determining how disparate data can be combined to reach a rational and scientifically supportable conclusion. In addition, EPA's weight-of-the-evidence process must:

- *Place greater emphasis on human studies.* Although EPA states that human studies (including epidemiological studies) should be given more weight than animal studies, in practice the Agency does not consistently follow this policy. In particular, EPA sometimes dismisses epidemiological studies of any quality that do not show positive associations and accepts with little resistance studies that yield positive associations irrespective of their scientific quality. Epidemiological studies of highly exposed occupational cohorts provide important information on the human toxicity of chemicals and should inform EPA toxicity assessments to a much larger extent than at present.

- *Use causation analysis.* Causation analysis, sometimes referred to as application of the Hill Criteria (Bradford-Hill 1966), should be used to evaluate whether exposure to a particular chemical may cause an increased risk of disease. Specifically, causation analysis should be applied to a group of studies that have investigated potential

associations between exposure to a particular chemical and a specific disease endpoint.

It bears emphasizing that EPA's 1996 Proposed Guidelines for Carcinogen Risk Assessment unequivocally endorse the weight-of-evidence approach for evaluating epidemiological data applicable to a particular chemical and describe "well-accepted criteria for causation" that should be used in such an approach. At least 10 cause-and-effect analytical criteria have been proposed, according to EPA's document, though only six are described as "fundamental" criteria (EPA 1996, pD-9). Section 6 of the Council's comments cites in full the relevant portions of the EPA document, which reads in part, "Analyzing the contribution of evidence from a body of human data requires examining available studies and weighing them in the context of well-accepted criteria for causation." The proposed guidelines, along with other EPA guidelines and statements, suggest that Agency risk assessors well understand the scientifically correct procedures to undertake a proper weight-of-the-evidence evaluation, even if the experience of regulated entities indicates EPA is not fully utilizing this understanding in practice.

⇒ *Agencies should accept site- or chemical-specific data.* Although EPA recommends use of site-or chemical-specific data, it often does not accept their use, requiring instead that conclusive or unambiguous evidence be provided before a default value can be superseded. OMB should direct agencies to use site- or chemical-specific information first, and if these data are unavailable, an agency may consider a safety or default value consistent with the above recommendation.

⇒ *Agencies should fully implement the Information Quality Guidelines.* OMB should insist that federal agencies fully apply their Information Quality Guidelines in the course of conducting risk assessments, and should do so in a manner that is consistent with OMB's government-wide standards. Agencies should defer to studies that meet these guidelines and must set aside potentially influential information that is not transparent enough to be reproducible, or data deemed to be of questionable utility or integrity. In addition, information quality and applicability must be the primary drivers for

weight-of-the- evidence procedures, causation analysis, and the use of site- or chemical specific data (see above).

Lastly, when appropriate – for instance, when a pesticide is engineered for deliberate toxicity or a potent chemical's widespread use may create broad exposure – OMB should recommend ecological risk assessments be conducted on effects at the population and/or community level rather than on individual receptors.

Generally, the compounded conservatism of EPA risk assessments, together with a lack of uncertainty analysis, lead to regulatory decisions by both EPA, and state agencies that follow EPA's direction, that are unjustified relative to circumstances. This problem is readily apparent when overestimations of risk are used to justify CERCLA remedial actions. For example, the cost of remediating a site to a 1 ppm action level may be substantially higher than attaining a 5 ppm action level. Yet, because the risk assessments compile multiple layers of conservatism, both action levels are likely associated with no risk to the potentially exposed population and the costs incurred in attaining the lower level do not have any commensurate risk benefit.

Although, as noted above, the application of uncertainty factors and default assumptions is well accepted in risk assessment as a means to deal with data gaps and uncertainties, EPA risk assessments have employed judgments to inflate risk estimates beyond what is realistically justified by the scientific evidence. The Council's comments provide specific examples of EPA actions that, to use OMB's term, present highly "unbalanced" risk assessments. As a longer-term matter, OMB should work with EPA to plan and implement a program for upgrading Agency risk methodologies to reflect the state-of-the-science. This is imperative because EPA's risk assessments have far-reaching impacts on the U.S. economy, institutions, and public understanding of risks and should not be allowed to continue contributing to the "paranoia and neglect" that Dr. John Graham has accurately characterized as the state of our nation's risk management policies. Overall, our experience, as well as the concerns raised over many years by academic researchers and others who follow environmental risk issues, clearly point to the need for dramatic improvements in EPA risk assessments, including a more genuine application of the Agency's own scientifically sound and reasonable guidelines where available, such as the Agency's policy on probabilistic analysis.

The Council appreciates OMB's interest and efforts in fostering accurate, balanced and cost-effective risk assessment and risk management. Although the Council fully understands that an appropriate degree of caution should be used in making risk management decisions, such caution should be applied transparently in the risk management phase, and not opaquely in the risk assessment phase. Risk assessors, for their part, must seek to provide the best possible and most objective estimates of risk, fully disclosing any default assumptions necessitated by data gaps and uncertainties. If these fundamental changes are actually implemented in by EPA, all sectors will benefit.

III. COMMENTS

A. Draft Guidelines for the Conduct of Regulatory Analysis and the Format of Accounting Statements

1. SUMMARY

Overall, the draft guidelines are well written, comprehensive, and strongly grounded in economic theory and practice. Several new parameters for regulatory analysis are proposed:

- Use of both benefit-cost analysis (BCA) and cost-effectiveness analysis (CEA) in evaluating proposed regulatory actions
- Use of both 3% and 7% discount rates, and a 1% discount rate for regulatory actions with intergenerational effects (in addition to 3% and 7%).
- Conduct of formal quantitative uncertainty analysis (e.g., Monte Carlo analysis) for rules having an impact of $1 billion or more

The Council believes that the conduct of both BCA and CEA will greatly strengthen regulatory analyses since each addresses a different question. BCA addresses the question of whether a proposed regulatory action is "worth it" (i.e., do the benefits exceed costs) while CEA addresses the question of regulatory "effectiveness" (e.g., cost per unit of benefit, such as “lives saved”). For discount rates, selection of the most appropriate rate is dependent on factors such as the economic status of the population segment impacted (more wealthy segments may have lower discount rates vs. poorer segments) and the outcome being valued (e.g., health improvements vs. non-health outcomes such as energy consumption reductions). As such, we strongly recommend that agencies should not limit their analyses to a couple of arbitrary rates that may or may not be appropriate. They should begin with the consistent 7% discount value as a “weak” default value and perform sensitivity analyses across a very wide swath of discount rates that could be reasonable under particular circumstances for specific subpopulations. OMB must require that when an agency expresses a preference for a specific discount rate, it must provide to OMB a cogent economic rationale and support it with specific data for that regulatory action. We believe that OMB must direct agencies to scale the degree of uncertainty analysis based on the

impact of the proposed regulation rather than setting a hard and fast threshold (i.e., $1 billion). Finally, we believe that these draft guidelines would benefit from more consistent and transparent language related to the need for sensitivity analysis. OMB must clearly and objectively discuss the limitations associated with value of statistical life (VSL) estimates derived from occupational wage/risk premium studies. Overall, we believe that with these enhancements the OMB guidelines will provide sound regulatory analysis guidance and improve the utility and quality of economic information for regulatory decision making.

2. ASSESSMENT OF DRAFT OMB REGULATORY ANALYSIS GUIDELINES

a) Strengths of the Guidelines

These guidelines[3] are well written and clear in intent. OMB provides cogent explanations for several important economic principles (e.g., opportunity cost, benefits transfer). As the proposed guidelines have a number of strengths, only the most notable are described here. The first is OMB's clear emphasis on the need for agencies to evaluate alternative regulatory actions. OMB appropriately devotes an entire section of the guidelines (Section II) to describing possible alternative actions that should be considered (e.g., different enforcement methods, different degrees of stringency). We believe that OMB should direct agencies to perform these analyses, as too often agencies perform economic analyses limited to their preferred regulatory action rather than presenting benefits and costs of all reasonable actions. Evaluating all alternatives provides decision makers with the information needed to identify the alternative that "maximizes societal net benefits". We also support OMB's guidance that agencies identify and consider the undesirable side effects and ancillary benefits possibly associated with the proposed regulatory action. Many analyses to date have provided seemingly little attention to systematically evaluating the unintended consequences of proposed regulatory actions. Finally, the Council supports OMB's direction to agencies that they seek opinions, early in the analysis process, of those who will be directly affected by the regulation. In particular, industry is likely to have important information to offer to strengthen assessments, such as a better understanding of the

[3] Office of Management and Budget (OMB) Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations, Appendix C: OMB Draft Guidelines for the Conduct of Regulatory Analysis and the Format of Accounting Statements. Federal Register 2003; 68(2):5492-5527.

opportunity costs (and not just *compliance* costs) of proposed regulatory action or potential unintended consequences. We believe that it is crucial for this information to be collected and appropriately used by the agencies in their regulatory analysis.

b) Areas for Improvement

We identified several areas where the guidelines could be strengthened. Sensitivity Analysis: In particular, OMB's discussion of the importance of sensitivity analysis is weak in spots. For example, OMB states that "it is *usually* helpful to provide a sensitivity analysis," but does not illustrate any situation when sensitivity analysis is *not* helpful.[4] OMB should take a much stronger stand in favor of routine performance of sensitivity analysis. An agency's failure to perform sensitivity analysis on parameters that have a material effect on net present value benefits should generally be interpreted as a serious analytic defect. Indeed, elsewhere OMB has stated called sensitivity analysis "an essential feature of high-quality analysis."[5]

VSL Estimation: There are only limited discussions regarding drawbacks associated with value of statistical life (VSL) estimates derived from occupational wage premium studies. Most VSL estimates used by agencies today are derived from studies of death rates by occupation and wage premiums associated with more "risky" jobs. A sentence on page 5519 (section IV a. 4, 2nd to last paragraph) describes two types of potential bias associated with these VSL estimates: differences in risk tolerance and differences in the "voluntariness" of risk in the study and populations affected by regulation. OMB should elaborate more on how the "voluntariness" of a risk can be objectively assessed, for the distinction between voluntary and involuntary risks is often more difficult to make in practice than it has been to describe in theory. Some risks that people often construe as voluntary have substantial involuntary components. Much of the risk associated with driving a car arises from the actions of others over whom we have no control.

---

[4] See page 5514, col. 2, penultimate paragraph (emphasis added).

5 "…[S]ensitivity analysis is widely regarded as an essential feature of high-quality analysis, yet sensitivity analysis cannot be undertaken by outside parties unless a high degree of transparency is achieved. The OMB [information quality] guidelines do not compel such sensitivity analysis as a necessary dimension of quality, but the transparency achieved by reproducibility will allow the public to undertake sensitivity studies of interest." See Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication, 67 FR 8456.

Some general cautions are presented about benefits transfer issues[6] associated with VSL estimates but nothing specific to occupational wage premium studies. Given the widespread use by agencies of VSL estimates from occupational wage premium studies, a more detailed discussion of the limitations of these studies is clearly warranted. One obvious problem is that the commodity valued in wage premium studies may be significantly different from the regulatory context to which an agency might want to apply it. Mortality valuations apply to much more complex commodities than they readily appear to.

In addition, OMB should provide a more thorough discussion concerning the accuracy of transferring VSL values from wage premium studies to regulatory contexts. Just as it is scientifically dubious to simply extract and apply risk estimates (such as IRIS values) without carefully considering the context, it is economically problematic to simply pull VSLs "off the shelf" without giving serious concern to whether the transfer is appropriate. For example, it has been routinely assumed that the risk premium embedded in wages corresponds to the *objectively estimated* level of the specific occupational risk in question and not what workers *perceived* its magnitude to be. VSLs are understated (overstated) if workers believe that the occupational risks leading to wage premiums are lower (higher) than they really are. Estimates from well-crafted hedonic models should not be applied without adjustment if actual and perceived risks are not the same. Agencies using VSLs derived from these studies need to carefully examine whether the VSLs were based on actual or perceived risk and adjust the VSL values they use accordingly.

The improper use of VSLs from wage premium studies poses potentially serious information quality problems. As OMB has elsewhere stated, the "objectivity" component of information quality applies to *both* original *and* supporting data. Regulatory Impact Analyses clearly qualify as supporting data, and the improper use of VSLs violates OMB's objectivity standard because it disseminates information that is inaccurate, unreliable and biased.[7]

---

[6] Benefits transfer is defined as the transfer of existing estimates of non-market values from the context of a study to a new [policy] context

7 [… `O]bjectivity" involves a focus on ensuring accurate, reliable, and unbiased information. In a scientific, financial, or statistical context, the original and supporting data shall be generated, and the analytic results shall be developed, using sound statistical and research methods. See §V.3.b, 67 FR 8459 (emphasis added).

Benefit-cost Analysis (BCA)/Cost-effectiveness Analysis (CEA): OMB states that agencies should conduct both benefit-cost analysis (BCA) and cost-effectiveness analysis (CEA) whenever possible. OMB's rationale is that both analyses have strengths and drawbacks and that the two techniques "offer regulators somewhat different but useful perspectives and more robust information about tradeoffs." In particular, the two analyses address different questions: BCA addresses the question of whether a proposed regulatory action is "worth it" (i.e., under what conditions do social benefits exceed social costs) while CEA addresses the question of regulatory "effectiveness" (e.g., cost [$] per life saved). As such, performance of both types of analysis will greatly strengthen agency regulatory analysis. We strongly support this approach and recommend that OMB require that agencies perform both types of analyses wherever it is feasible to do so, and make this requirement a standardized “robustness check” for RIAs. In cases where agencies fail to conduct both BCA and CEA, OMB should require that they document why it is infeasible to do so. Making both analytical approaches normal requirements will give OMB (and the public) important additional information to determine if the agency has appropriately implemented these guidelines.

Discount Rates: The guidelines indicate that agencies should use both 3% and 7% discount rates, as well as a 1% discount rate for regulatory actions with intergenerational effects (in addition to 3% and 7%). The concept of discounting is based on the knowledge that people prefer present to future consumption and later to sooner cost-bearing. Thus, future benefits and costs must be discounted—and at the same rate. This is true irrespective of whether the benefits and costs in question are realized in financial terms, health effects, or other units.

Individuals may “discount” benefits or costs further when their realization is uncertain. Just as uncertain future benefits have less value, uncertain future costs are less worrisome. This is not discounting per se but the impact of risk aversion. In this case, the commodity can be characterized probabilistically as a likelihood function of a specific event, where willingness to pay is lower for probabilities of less than unity. OMB should make clear that matters of risk aversion should not be intermingled with the choice of the proper discount rate, which speaks only to the rate of time preference in consumption.

In addition, individuals may "discount" future health gains more than they "discount" financial gains. This phenomenon may arise because of doubts about one's future ability to actually enjoy health gains. The extent to which a certain future health gain event will yield today's estimated level of utility is conditioned on future health status, which is uncertain. OMB should emphasize that uncertainty over the attributes of the future-realized commodity also affects valuation but is unrelated to the choice of discount rate. Agencies should not confuse or intermingle commodity uncertainty with rates of time preference.

After ensuring that they have not commingled various forms of uncertainty with time preference, agencies should then carefully examine the regulatory outcome being valued (e.g., health improvements; non-health benefits, such as energy consumption reductions) and attributes of the affected population (e.g., healthy adults, children; the elderly or infirm; wealthy or poor households, taxpayers) for insights into the choice of discount rate. Many of these attributes influence the determination of the appropriate discount rates. For example, lower discount rates are more likely to apply to children than the elderly and the infirm; and to wealthy households more than poor households. Where regulatory beneficiaries are the same as regulatory burden-bearers, the most appropriate discount rate must consider these factors. The choice of discount rate implies the prior identification of the attributes of the affected population.

Distributional issues arise, however, where regulatory beneficiaries are demonstrably different from regulatory burden-bearers. In this case, there is no value-neutral principle that can be used to determine which discount rate is "best." A "low" discount rate favors those who are on average younger, healthier or richer by embedding their likely preferences into the analysis. Conversely, a "high" discount rate favors those who on average are older, ailing or poorer. A regulatory program that is intended to make older, ailing or poorer household better off would fail to cost-effectively achieve its objective if alternatives are analyzed and compared using discount rates lower than those held by the target beneficiary subpopulation. Indeed, using an artificially low discount rate could easily result in selecting an alternative that fails to benefit the target beneficiary subpopulation at all.

The fact that distributional differences are common between regulatory beneficiaries and regulatory burden-bearers provides yet another rationale for agencies to perform a vibrant

sensitivity analysis showing how net benefits vary depending on the choice of discount rate. This analysis provides decision-makers and the public critical information concerning how dependent net benefits are on the selection of discount rate and the location of the "break even" discount rate (i.e., the rate where present value benefits equal present value costs). We strongly recommend that OMB establish a 7% real discount rate as a “weak” default intended to permit ready comparison, but require that agencies perform sensitivity analyses across a very wide swath of discount rates that could be reasonable under particular circumstances for specific subpopulations.

We believe that this proposed approach is the proper one for three reasons. First, it draws attention to the fact that discount rates vary in the population and that no single discount rate applies to all. The details of a specific regulatory action are important factors in determining which rate is most reliable. Second, it avoids the temptation for agencies to bury this potentially critical policy-relevant factor as an “economic policy” default assumption. As we address in other sections of our comments, agencies establish “ policy” default assumptions or uncertainty factors that frequently engender tremendous controversy. The choice of discount rate is potentially another version of the same phenomenon. Third, our recommended approach reminds us all that the “right” discount rate may be a matter of genuine disagreement or policy controversy—especially in any case where regulatory beneficiaries are likely to have different rates of time preference than regulatory burden-bearers.

Where agencies express a preference for a particular discount rate, OMB’s guidance must require the agency to provide a cogent rationale that is supported by specific theory and relevant empirical evidence that specifically applies to that regulatory action. Off-the-shelf defaults are inherently suspect, as are rhetorical arguments, hypotheticals, or unrelated precedents. This approach ensures consistency across agencies while permitting substantial flexibility (as long as it is exercised rigorously), and has great utility for setting priorities across government programs.

Discounting presents a clear case where existing variation across (and sometimes within) agencies undermines the public’s capacity to make useful comparisons. OMB must set a high threshold for overcoming this consistent approach. The ability to perform programmatic comparisons is fundamental to the statute underlying this Report to Congress (“Regulatory Right

to-Know-Act") and is irreversibly harmed when agencies use inconsistent methods. Without consistency in discounting, potential health, environmental and safety risks cannot be assessed, compared and managed in an efficient and equitable fashion. Both the Administration and Congress lose their ability to prioritize programs and develop effective solutions.

Finally, with regard to a 1% discount rate for intergenerational effects, OMB provides no indication as to what may or may not constitute an "intergenerational" effect and thus threatens to undermine all that it accomplishes elsewhere with respect to discounting. *All* regulatory actions have intergenerational costs, as decisions made today reflect irreversible commitments of current and future resources and implicit decisions concerning what benefits to forego. In *some* cases, direct benefits are also intergenerational. It seems peculiar and seriously misguided to use completely different analytic principles for that small subset of cases where *benefits alone* are delayed by decades or centuries.

Apart from these concerns, there is little empirical evidence in the literature to support the choice of 1% over any other value. It appears that this special 1% rate is intended to apply to specific actions that are highly unlikely to yield net benefits under conventional methods. Further, this may embed controversial risk management preferences into what is supposed to be an objective analytic endeavor. We believe that this exception for so-called "intergenerational" effects must be removed from the guidelines and OMB should stick to a 7% "weak" default value, applied consistently across agencies, with additional requirements for robust sensitivity analysis of the effects of alternative discount rates across a very wide swath of discount rates that could be reasonable under particular circumstances for specific subpopulations.

Uncertainty Analysis: OMB indicates that agencies should perform formal quantitative uncertainty analysis (e.g., Monte Carlo analysis) for rules having costs of $1 billion or more. Clearly, uncertainty analyses will provide valuable information on the probability of occurrence (e.g., probabilities of harm to human health and safety) and aid decision-makers in determining whether to act now or seek additional information. However, defining any specific threshold will create incentives for agencies to avoid analyses by estimating impacts below the specified threshold, just as occurs today with OMB's $100 million threshold for economically significant regulations. More importantly, it makes no sense to establish any fixed threshold that triggers the

need for uncertainty analysis. The point of uncertainty analysis is to ascertain the expected likelihood that any fixed value will materialize. Equally disturbing is the presumption implicit in a fixed threshold that only uncertainty on the benefits side is important. Benefits assessments are already analytically richer than cost assessments; as OMB and others have frequently noted, agencies rarely, if ever, estimate opportunity costs. This is especially ironic because opportunity costs, which are benefits foregone, are so much more uncertain than direct benefits.

We believe OMB must establish a tiered approach and require that agencies scale the degree of uncertainty analysis to the level of impact of the proposed regulation. For example, for regulations having impacts of $1 billion or more a formal uncertainty analysis must be conducted, be exhaustive in scope and detail, and apply to both benefits and costs. Uncertainty analysis on the most critical parameters affecting net present value benefits ought to be performed for all economically significant regulations. For regulations with lesser impacts on the economy ($10 million to $100 million), the level of detail and scope of uncertainty analysis could be adjusted accordingly. In particular, OMB should seek uncertainty analysis on key regulatory determinants and model parameters, such as the uncertainty associated with projected costs for a regulatory action (to assess how likely it is that these effects will be unexpectedly significant). Armed with some information on the uncertainties on both the benefit and cost assessments, OMB can work with agencies to determine what additional information could be collected to improve and further inform agency decisions.

B. U.S. Government's Approaches to Analysis and Management of Emerging Risks

OMB has requested comments on current U.S. approaches to analysis and management of emerging risks. Specifically, OMB has requested comment on the following issues:

- Ways in which "precaution" is embedded in current risk assessment procedures through "conservative" assumptions in estimation of risk, or through explicit "protective" measures in management decisions as required by statutory requirements as well as agency judgments.

- Examples of approaches in human and ecological risk assessment and management methods addressed by U.S. regulatory agencies (e.g., consumer product safety, drug

approval, pesticide registration, protection of endangered species) that appear unbalanced.

The comments that follow address these issues by primarily focusing on EPA risk assessment practices. We first address how "precaution" should be factored into EPA risk assessment. Relying on important work by the AEI-Brookings Joint Center for Regulatory Studies[8], we next focus on the extent to which EPA environmental decision-making is, in fact, based on objective, realistic, and scientifically balanced risk assessments. We then address the extent to which EPA risk assessments overestimate risk, concluding that typical EPA risk assessments overstate risk by a factor of ten and, in some cases, perhaps by a factor of 100. We then provide some comments on the societal costs of EPA's failure to appreciate the imprecision of its risk estimates when using such estimates to make remedial or regulatory decisions. Finally, we offer some thoughts on how EPA risk assessment practices can be improved to more accurately estimate risk and reduce costs to society. In an appendix, we provide examples of cases where EPA risk assessments have grossly overstated risks.

1. Use of "Precaution" in Risk Assessment

The Council supports the intended approach of the U.S. regulatory system to rely on objective, realistic, and scientifically balanced risk assessment, to separate risk assessment from risk management, and to emphasize cost-benefit analysis, as spelled out in E.O. 12866. In the Council's view, the general risk assessment/risk management model as described by OMB and federal agency guidelines can, if properly implemented, provide for appropriate precaution in the risk management process without confusing the science-based risk assessment process with policy judgments relating to precaution. In instances where the potential costs to society are high, it may be appropriate to take a more cautionary approach in the measures taken to address clearly identified risks. Indeed, the Council believes that many regulatory measures already in place embody this approach.

---

[8] See Appendix 1 [W. Kip Viscusi and James T. Hamilton. April 1999. Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions (Working Paper 99-2). (Research support was provided under Cooperative agreement No. CR-823604-01 from the U.S. EPA Office of Policy, Planning, and Evaluation.)].

The Council also firmly believes that industry and government must work together to continually enhance the data used in risk assessments, with the goal of eliminating or reducing the need for default assumptions in the risk assessment process to the maximum extent possible. Using overly restrictive precautionary measures can deprive society of important benefits to human health, environmental quality, and improvements in the quality of life. The Council supports an approach that provides for a reasonable and cost-effective response whenever a risk assessment yields reasonable evidence of the risk of serious or irreversible harm to health or the environment because of a particular product or activity.

In the Council's experience, however, government agencies, including EPA, have gone well beyond a science-based approach in assessing and managing perceived risks. As detailed in Appendix 5 of this document, there have been highly unbalanced government interpretations of precaution that essentially reject sound science and hamper innovation, instead favoring exaggerated risk assessments as the basis for precautionary action.[9] Current EPA assessment practices and policies often have disregarded real-life data and site-specific information in favor of multiple layers of overly conservative default assumptions. The Council opposes decision-making processes that do not have a strong basis in science and objectivity. Besides citing examples in Appendix 5 of specific rules and decisions based on unbalanced applications of safety or uncertainty factors, the Council's comments also list in Appendix 4 the broad range of conservative default assumptions available through EPA methodologies, guidelines, and practices.

[9] Perhaps the most egregious example stems from a decision by Peruvian officials the 1980s. Based on USEPA studies, Peruvian officials wanted to reduce cancer risks associated with trihalomethanes ("THMs") in drinking water. To achieve this goal, these officials and others from neighboring Latin American countries abandoned chlorination of their drinking water supplies, thereby contributing to a cholera epidemic that killed at least 3,500 people. In response, the Pan American Health Organization sent a letter to then-USEPA Administrator William Reilly asking for a letter clarifying that chlorination to control waterborne diseases "should be afforded top priority." A World Health Organization official editorialized that the uncertainties associated with the THM cancer risk assessment should have been balanced against the "disaster potential of not disinfecting water supplies."

2. Whether Risk Truly Drives EPA Decision-making

Environmental regulation in the United States is based on a combination of technology-based controls and risk assessment. For example, the regulation of pollutant discharges to waters of the United States is based on compliance with technology-based standards, with residual risks being protected against by water quality-based effluent limitations. Under other statutes, like CERCLA, remedial standards are to be based primarily on risk; technological considerations may be relevant after it is determined that an unacceptable risk exists (e.g., if a landfill is deemed to pose an unacceptable risk, a cap meeting RCRA design standards might be installed).

Although science-based risk is intended as the basis of U.S. environmental regulation, scholars have questioned whether risk truly drives CERCLA remedial requirements imposed by EPA at various sites. In 1999, on behalf of the AEI-Brookings Joint Center for Regulatory Studies, W. Kip Viscusi of Harvard Law School and James T. Hamilton of Duke University's Sanford Institute of Public Policy used original data on the cleanup of 267 hazardous waste sites to examine whether factors other than risk influence remediation decisions [Viscusi and Hamilton, 1999 (Appendix 1)]. The authors collected cost information on all 267 sites and risk data on a subsample of 150 sites. This yielded a human health risk database with information on over 20,000 chemical risk pathways at the 150 sites, which enabled the authors to develop estimates of the number of cancer cases averted by remediation and the cost per case of cancer averted. In general, Viscusi and Hamilton found that, although decisions under CERCLA are supposed to be based on risk, other factors – such as economics and politics – seemed to bear more heavily on remedy selection than risk reduction

The authors found that although EPA regulation and guidance provides the decision maker discretion regarding whether remedial actions will be taken at sites with cancer risks between 10^{-4} and 10^{-6} – and also provides that remedial actions should achieve risk level within this same range – in practice the cleanup goal chosen by EPA is often more stringent than 10^{-6}. The authors attributed this finding to the fact that EPA guidance encourages conservatism in exposure scenarios (e.g., future residential land use is often assumed even if the surrounding area is industrial) and parameter assumptions (e.g., the 95 percent confidence limit on the estimate of

the mean concentration of the chemical is often used to represent a chemical's concentration at a site).

Viscusi and Hamilton concluded that, overall, CERCLA expenditures were not cost-effective when evaluated in terms of cancer prevention. The mean number of cancer cases averted per site over a 30-year period was 5.6, with a range from 0 to 652 and a median of 0.019. The authors reported that the mean cost per case of cancer averted was $11.7 billion. The median cost was $418 million. The range was from less than $20,000 to $961 billion. At only 36 of the 150 sites was the cost below $100 million per cancer case averted. These estimates use EPA conservative risk assumptions and assume no latency period. With adjustments for these factors, the median cost rises to above $1 billion per cancer case. The most effective 5% of all cleanup expenditures eliminated over 99% of the cancer risk. Stated otherwise, 95% of the costs are spent to address less than 1% of the risks.

Finally, Viscusi and Hamilton found it disturbing that the presence of actual risk to people based on current land use patterns did not increase the stringency of site cleanup goals. That is, pathways exposing current residents generally did not receive more stringent standards than pathways that might expose hypothetical future uses. The authors therefore concluded that EPA is failing to target its efforts to protect currently exposed populations. The authors concluded their analysis as follows:

> *EPA cleanup policies are an outlier among government regulatory programs on any efficiency basis, assuming cancer prevention is the primary objective. The benefits of Superfund cleanup are highly concentrated at a very small percentage of sites, with most cleanup actions failing any reasonable efficiency test. The . . . results highlighted the pivotal role of political factors for inefficient cleanups, whereas the most desirable cleanups were not influenced by voting rates.*

The findings of Viscusi and Hamilton lead to the important question of whether EPA risk assessment practices are the cause of the problems many external parties perceive in EPA risk-based regulatory and remedial programs. If, as the authors suggest, risk assessment is being used to justify decisions made for other reasons, and not as the basis for scientifically well-founded

regulatory and remedial decisions, only by raising the quality of EPA risk assessments can the problem be adequately addressed.

3. EPA Does Not Accurately Estimate Risk

There can be little doubt that EPA risk assessors deal with uncertainty in assessing risk by systematically overestimating risk. Indeed, EPA has explained its process as follows:

> *To account for these uncertainties and to acknowledge gaps in science, we build in safety factors in the risk estimates which tend to overestimate what we believe to be the actual risk. Where there is uncertainty or where our information is incomplete, we make assumptions that tend to overestimate the risks as a way to insure the public health is protected. . . . As a result, when we estimate that there is a one-in-one-million (excess) risk, the actual excess risk is probably much less and may even be zero.*

EPA (2000)

While it is a well-accepted risk assessment practice to use default assumptions in the face of uncertainties about toxicity and exposures, this practice should not be taken as a license to insert multiple layers of conservative assumptions that in the name of uncertainty distort scientific evidence so that assessments cease to bear any resemblance to objective, realistic, and scientifically balanced evaluations of the data. Understanding that EPA risk assessment intentionally err on the side of overestimating risks, it is natural to seek to determine the extent to which that is true. A few examples are sufficient to illustrate, qualitatively, that EPA risk assessments result in very large overstatements of risk:

In deriving quantitative estimates of chemical toxicity through the calculation of oral Reference Doses ("RfDs") or inhalation Reference Concentrations ("RfCs"), EPA divides measures of toxicity from animal or human studies to account for up to five possible causes of uncertainty. By convention, the uncertainty factors range from 1 to 10 and are usually combined by multiplication, not addition. The process of dividing toxicity estimates by uncertainty factors means that all issues of uncertainty are resolved conservatively. For example, where data from an animal study are used to estimate toxicity to humans, one uncertainty factor – often a factor of 10 – is used to account for the possibility that humans are more sensitive than the test species.

So, if the animal species and humans are equally sensitive to the chemical, this one factor alone assures that all risk assessments using the RfD will overestimate risk by a factor of 10. If the animal species is more sensitive than humans, risk assessments using the RfD will overestimate risk by more than a factor of 10.

Typically, the products of the several uncertainty factors used in deriving RfDs or RfCs are very large. For 213 out of 414 chemicals (51%), the RfDs or RfCs in EPA's Integrated Risk Information System (IRIS) database were derived using total uncertainty factors greater than 100 (IRIS, 2003). For example, in the case of EPA's RfD for PCB Aroclor 1254, four uncertainty factors were multiplied together and divided into a toxicity measurement from a Rhesus monkey study to calculate the human RfD: (1) a factor of three was used to account for the possibility that humans are more sensitive than monkeys; (2) a factor of ten was used to account for the possibility that some humans are ultrasensitive to PCBs; (3) a factor of three was used because the monkey study was not a full lifetime study; and (4) a factor of three was used because the toxicity measurement from the monkey study was a "lowest observed adverse effect level" rather than a "no observed adverse effect level." The total uncertainty factor was 300 (rounded up from 270, the product of the four uncertainty factors). As discussed in AMEC (2002) (Appendix 2), the evidence is that Rhesus monkeys are more sensitive than humans (not the other way around), that the "less than lifetime" uncertainty factor is unnecessary, and that an "interindividual sensitivity" factor of three is more than sufficient. Thus, the total uncertainty factor of 300 is clearly excessive.

In performing risk assessments for possible carcinogens, EPA uses "cancer slope factors" ("CSFs") to estimate toxicity.[10] CSFs are typically derived from high-dose animal study tumor frequency data using a hypothetical dose-response curve. Moreover, the CSF is taken from the upper bound slope of the hypothetical dose-response curve. Nearly 20 years ago EPA acknowledged that this approach overstates risk:

---

[10] A CSF is the upper bound, approximating a 95% confidence limit, on the increased cancer risk from a lifetime of exposure to an agent. This estimate, usually expressed in units of proportion (of a population) affected per mg/kg/day, is generally reserved for use in the low-dose region of the dose-response relationship, that is, for exposures corresponding to risks less than 1 in 100. See www.epa.gov/iris/gloss8.htm.

> *It should be emphasized that the linearized multistage procedure leads to a plausible upper limit to the risk that is consistent with some mechanism of carcinogenesis. Such an estimate, however, does not necessarily give a realistic prediction of the risk. The true value of the risk is unknown and may be as low as zero.*

EPA (1986).

The factors EPA uses to estimate chemical exposures are biased. EPA's "high-end" exposure assumptions are so conservative that the individuals EPA is seeking to protect could not reasonably be expected to exist. For example, in seeking to protect people from chemicals in soil, EPA imagines an individual who eats 200 milligrams ("mg") of soil every day as a child, then eats 100 mg every day for 24 or more years as a adult from the same area, breathes in dust from the same place every day, and has soil stick to his or her skin for an extended period of time over a large area of this person's body – again every day. If this person then grows up to be a construction worker, EPA assumes, without any supporting data, that the person eats 480 mg of soil per day on the job. This person is truly imaginary.

The conservatism of EPA's exposure assumptions is compounded by the fact that when an exposure assessment uses a number of data points (or assumptions) that are at or near their maxima (i.e., 95th percentile values), the resulting exposure estimate represents a condition which rarely, if ever, could occur. This is illustrated by Burmaster and Lehr (1991), in which the authors provide a simple relationship from probability theory that describes the likelihood of occurrence of an outcome based on a series of conservative assumptions. If, for example, three conservative (95th percentile) exposure estimates are multiplied together, the outcome actually represents the 99.99th percentile of exposure (not the 95th percentile exposure), based on the following equation:

$$1 - (1 - 0.95)^3 = 0.9999$$

In other words, only 0.01% of a given population (1 in ten thousand, or 10 in 100,000, or 100 in 1,000,000) would experience exposure at or greater than this level. Thus, if – for example – an EPA baseline risk assessment for a site predicted a 1×10^{-4} risk for the maximally exposed individual living near the site, and 10,000 people live near the site, then 0.0001 extra people (one one-thousandth of a person) would be expected to contract cancer from living near the site.

It can thus be seen that EPA risk assessments inaccurately estimate actual risk. But can one determine the extent to which EPA risk assessments overstate risks? Although we believe it is difficult to answer this question with a high degree of precision, as discussed below it is possible to make an order of magnitude estimate of the extent of risk overstatement inherent in EPA's risk assessment approach.

4. Societal Cost of EPA's Failure to Provide Reliable Risk Estimates

The fact that EPA risk estimates are highly biased is exacerbated by the manner in which risk estimates are typically expressed by the Agency. The Agency's risk assessment approach is mathematical – numbers are added, multiplied and divided in the risk assessment calculation and the risk estimate itself is presented as a number. Because of this mathematical approach, the general public and many regulators believe, contrary to the fact, that the risk assessment process produces risk estimates that are reasonably precise. EOP (1991). In other words, there appears to be a general belief that the outputs of risk assessment calculations can be used to predict the actual potential for injury or disease resulting from chemical exposure. EPA fosters this perception by the precision with which it presents risk estimates. EPA typically uses two significant figures in cancer estimates – e.g., "3.2×10^{-5}" – rather than simply expressing the risk as "less than 1 in 10,000."

The imbalance in EPA risk assessment practices, together with the suggestion of precision where little may exist, can have substantial economic significance because regulatory and remedial decisions are commonly made by EPA based on the incorrect belief that differences in risk estimates reflect actual benefits or detriments to human health or the environment. This problem is readily apparent when admitted overestimations of risk are translated into, or used as justification for, CERCLA remedial actions. For example, the cost associated with cleaning a site to a 1 ppm action level may be substantially higher than remediating the site to 5 ppm. Because of the misconception or misinterpretation of the precision of the risk assessment process, regulators often believe that there is an actual benefit to human health or the environment in attaining the lower remedial objective and, conversely, that there is an unacceptable level of risk reduction in attaining only the higher value. The fact is that, because of the highly conservative bias in Agency risk assessments, both of these action levels are in all

probability associated with no risk to the potentially exposed population and the costs incurred in attaining the lower level do not have any commensurate risk-reduction benefit.

EPA risk assessors may well respond to the criticism that the Agency fails to appreciate the imprecision of its risk assessments by pointing out that virtually all of its risk assessments conclude with a qualitative (not quantitative) discussion of uncertainty. Although that is true, it obscures the fact that EPA *risk management decisions* typically do not reflect the uncertainty in the assessment. Rather, EPA risk managers engage in discussions regarding whether a cleanup or other regulatory standard (e.g., water quality standard or maximum contaminant level) should be 1 ppm or 2 ppm without acknowledging that it likely makes no significant difference in risk (although it may in cost). The Council recommends that this issue be given high priority by OMB and that efforts be made to assure that EPA, as well as other federal agencies, make regulatory decisions that reflect the substantial uncertainty inherent in risk assessment.

5. Examples of EPA Risk Assessments that Inaccurately Estimate Risks

In this section we provide some examples (summarized from the fuller discussion found in Appendix 5) of EPA decisions made during Agency risk assessments that have resulted in substantial and potentially very costly biased estimates of risk. EPA decisions leading to overstatement of actual risks have been made in connection with all aspects of the human health and ecological risk assessment processes, including environmental sampling and data collection, exposure assessment, toxicity assessment and risk characterization. Both the Table 1 below and Appendix 5, containing the details, are divided into two major sections, the first of which discusses human health risk assessment and the second of which discusses ecological risk assessment. Within each major section, examples are organized by major steps in the risk assessment process.

TABLE 1 -- EXAMPLES OF EPA INACCURATE RISK ASSESSMENTS	
Human Health Risk Assessment ("HHRA")	
Exposure Assessment	
Name	*Summary*
Gas Turbine Ass'n	• EPA requiring that human health risk assessment ("HHRA") use

petition to "delist" gas turbines from MACT	exposure period of 70 years, rather than the conventional 30 years for this type of assessment. · EPA requiring that HHRA assume that maximally exposed individual never leaves home. · EPA requiring that HHRA use highest possible emission factor for gas turbines.
Fox River HHRA	· EPA deriving high, long-term, fish consumption estimate from study that did not measure long-term consumption instead of from study that did measure long-term consumption. · EPA using out-of-date chemical concentration data that do not reflect current exposure. · EPA assuming high consumption of fish species (carp) that is rarely eaten. · EPA using exposure periods that are about two times too long based on available data.
Hudson River HHRA	EPA's HHRA for the Hudson River overestimates human health risk because it grossly overstates the rate at which Hudson anglers consume fish from the River. Because EPA failed to conduct a Hudson angler survey, it needed to rely on studies of other water bodies. Although there are five studies that could be used to estimate fish consumption, EPA relied on a single study (Connelly et al. 1992) that reported consumption rates four times higher than the average of other, better conducted, studies. EPA's use of the 1992 Connelly study was inappropriate for several reasons. · EPA derived a consumption rate almost three times greater than the authors of the study found. · The study was not designed to assess consumption rates, but rather angler awareness of and knowledge about fish consumption advisories; as a result, numerous assumptions were required to generate consumption rates. · Individuals who do not respond to surveys of this type are likely to consume considerably less fish than individuals who do respond. The response rate reported by Connelly is on the low-end of acceptable standards, which biases fish consumption estimates toward higher level consumers · The consumption rates based on Connelly et al. (1992) are inconsistent with well-conducted studies of similar angler populations

	which are more appropriate for estimating rates of fish consumption for the Hudson.
Housatonic River HHRA	· EPA assuming excessive exposure frequency and duration (e.g., 61 days/yr for trespasser in inaccessible area for period of nine years). · EPA assuming excessive dermal contact (up to six square feet of skin covered with soil on each visit to site). · EPA assuming excessive soil ingestion (up to 100 milligrams of soil on each visit to the site). · Such assumptions used even though blood concentrations of nearby residents had been measured and were not elevated.
Manistique Harbor HHRA	The Baseline Human Health Risk Assessment (BHHRA) for Manistique Harbor calculated a cancer risk to the average and "high-end" recreational angler of 1.8×10^{-5} and 2.4×10^{-3}, respectively. For the "average" and "high-end" subsistence anglers the risks were 2×10^{-4} and 1.2×10^{-2}, respectively. These values were based on several overly conservative assumptions, including: · The high-end angler scenarios assumed 25% of the fish diet was carp, despite the finding that few if any Upper Peninsula anglers regularly consume carp. · It was assumed that subsistence anglers obtained 50% or 100% of their fish from Manistique Harbor. This was unlikely given the demographics of the population and the difficulty associated with fishing from the banks of the Harbor. Manistique Harbor is small, the banks are bulkheaded, and better and more accessible fishing areas on Lake Michigan are readily available. · It was also assumed that the anglers consumed fish from the Harbor 365 days a year. Since the Harbor freezes over in the winter, this assumption is wholly unfounded.
Change in adult default soil ingestion date	· Even though in 1997 EPA relied on a 1990 study to establish a default adult soil ingestion rate of 50 mg/day, in 2001, without explanation, EPA relied on the same study to change the default value to 100 mg/day. · EPA's has no scientific basis to change default value to 100 mg/day, especially since a 1997 study supports a default adult soil ingestion value of 20 – 40 mg/day.
Refusal to alter construction worker	EPA continues to use default construction worker soil consumption rate of 480 mg/day based on the conjectures found in Hawley (1985) even

soil consumption rate	though: • New measurements of soil to skin adherence show that the Hawley (1985) assumptions were excessive. • Using average and high-end soil adherence rates developed by an EPA workgroup, the construction worker soil ingestion rate ranges from 33 mg/day (average) to 64 mg/day (high-end).
EPA continues to use excessive dermal absorption factor for PCBs	EPA's dermal absorption factor for PCBs of 14% overestimates the fraction of PCBs that are absorbed through the skin because: • Study on which it is based used soil with very low carbon content. • Study methodology did not mimic expected chemical mixtures (fresh vs. weathered PCBs) or conditions of dermal exposure (24-hour monkey exposure through abdominal skin vs. shorter-term human exposure through hands) • Recent study using soil with 5-6% carbon content, weathered PCBs, and 12-hour dermal exposure period supports dermal absorption of approximately 4%.
EPA failure to perform probabilistic risk assessments	Although EPA guidance endorses the use of probabilistic risk assessment methodologies and has published detailed guidance regarding the topic, EPA ignores its own guidance. For examples: • The design of EPA's probabilistic model for the Hudson River was seriously flawed. The design forced EPA to assume that anglers consumed unrealistic amounts of fish harvested from the same locations, cooked in the same fashion, and composed of the same mixture of species every year for more than 30 years. The model did not account for variation in human behavior nor did it account for declining concentrations in fish tissue contaminant levels over time. • EPA Region 5 and Wisconsin Department of Natural Resources ("WDNR") conducted a HHRA that relied principally on a "point estimate" or deterministic approach in arriving at estimates of cancer and noncancer risk from consuming fish from the Fox River. Although the EPA/WDNR HHRA used certain probabilistic methods as part of a sensitivity analysis, the HHRA did not include a probabilistic risk assessment. The combination of incorrect input parameters for fish consumption, fish tissue concentrations, and population mobility, together with the multiplicative nature of deterministic risk assessments assured that the Fox River HHRA overestimated risk by up to several orders of magnitude.

Toxicity Assessment	
Name	*Summary*
EPA cancer risk assessment procedures	⬩ EPA generally requires use of linearized multistage ("LMS") model to estimate risk even though EPA knows the model "does not necessarily give a realistic prediction of the risk" and that "[t]he true value of the risk is unknown, and may be as low as zero." ⬩ EPA compounds the conservatism of the LMS by using the lower 95% limit on the dose that is estimated to cause a 10% response to derive a CSF, rather than a more central measure of dose. ⬩ Use of a model other than the LMS is permissible only if stringent conditions – which can rarely be met – are satisfied. ⬩ EPA will rarely give any weight to negative results in human cancer studies. ⬩ If human data are unavailable, EPA's default is that positive effects in animal studies mean the chemical may cause cancer in humans. Combined with preceding principle, this means that a single positive rat study may trump several negative human studies. ⬩ Another EPA default is that "effects seen at the highest dose . . . are appropriate for assessment," even though EPA knows that using maximum tolerated dose to project low dose effects is highly questionable. ⬩ A final EPA default is that "target organ concordance is not a prerequisite for evaluating the implications of animal study results for humans." This allows EPA to assume that a chemical is a human carcinogen even when there is evidence that humans do not contract the same type of cancer as test animals.
EPA's increase in IUR for 1,3-butadiene	In deriving new IUR for 1,3-butadiene, EPA: ⬩ Ignored SAB advice and excluded high exposure individuals from dose-response modeling, thus inflating the cancer estimate. ⬩ Departed from its practice of using the maximum likelihood estimate and instead used the 95% upper confidence limit, thus inflating the cancer estimate. ⬩ Ignored SAB advice to model risk using the window of exposure model, and instead used cumulative lifetime exposure, thereby inflating the risk estimate.

	• Without justification, used a longer than standard exposure duration (85 yrs rather than 70 yrs). • Without explanation, departed from its practice of estimating lifetime cancer risk using mortality rates (rather than incidence rates). • Contrary to guidance and accepted practice, and without scientific justification, applied a "gender uncertainty factor" in deriving the IUR. • Failed to give consideration to the cumulative impact of its many "health protective" choices, resulting in an IUR 20-fold more stringent than it proposed in 1999.
RfD for perchlorate	In January, 2002, EPA published a toxicity assessment for perchlorate, an anion that mimics iodide and may effect thyroid hormone levels. EPA recommended an RfD of 0.00003 mg/kg-day This value equates to a drinking water level of 1 ppb. There is no supportable scientific basis for the draft perchlorate RfD because: (1) the RfD is based on a NOAEL from highly suspect rodent data and application of an uncertainty factor of 300; and (2) the human data indicate that perchlorate is not toxic at levels at least 200 times higher than EPA's RfD. That evidence includes: • Perchlorate has been used as a medication to treat hyperthyroidism associated with Grave's disease. Adult dosages of potassium perchlorate of 200 – 900 mg/day produce clinical results. • Two human studies indicate no adverse thyroid or other health effects at perchlorate dosages up to 0.7 mg/kg-day. • A human volunteer study with 10mg/day perchlorate dosing for two weeks showed no changes to thyroid hormone levels. • Another human volunteer study, which EPA helped design, involved doses ranging from an equivalent of 200 ppb to 17,000 ppb perchlorate in water. No hormone effects were observed at the high dose. • Perchlorate occurs naturally in northern Chile. There were no adverse thyroid or any other health differences attributable to life long exposure to perchlorate at 110 ppb. • No differences in neonatal thyroid hormone levels or Medicaid data regarding prevalence of thyroid diseases or cancer were found in exposed and non-exposed infants from Las Vegas and Reno, Nevada, respectively. • There is no increase in neonatal hypothyroidism in southern

	California in zip codes associated with elevated perchlorate exposure. Based on the human studies and appropriate uncertainty factors, the RfD for perchlorate should be 0.005 to 0.17 milligrams per kilogram of body weight per day, equivalent to 175 to 6,000 parts per billion in drinking water.
Proposed RfD for acetone	EPA proposed an RfD for acetone of 0.3 mg/kg/day. This value is scientifically unsupportable because: · It is more than 100-fold below normal endogenous production of acetone in healthy individuals. Thus, a daily dosage of the magnitude of the RfD is meaningless from a toxicological perspective. · It is inconsistent with toxicity assessments performed by several other scientists and groups, including WHO, which has published a recommended value of 9.0 mg/kg/day, 30-fold higher than EPA's RfD · Acetone exhibited very low toxicity in 90-day drinking water studies sponsored by the National Toxicology Program ("NTP"). Minimally toxic concentrations were estimated to be 20,000 ppm or higher for males and females of different rodent species (20,000 ppm ≈ 1,700 mg/kg/day for male rats). NTP recommended against conduct of chronic studies because "the prechronic studies only demonstrated a very mild toxic response at very high doses in rodents." · In deriving the proposed RfD, EPA applied a combined total uncertainty factor of 3000, which is demonstrably overconservative.
Proposed RfD for trichloroethylene	To derive the RfD, EPA used an uncertainty factor ("UF") of 50 for human variation and values of 3 each for animal to human extrapolation, subchronic to chronic exposure, and LOAEL to NOAEL extrapolation. A "modifying factor" of 3 was also applied to reflect background exposure. This resulted in an overall UF of 5000, which was lowered to 3000, EPA's maximum UF. This UF is unnecessarily stringent because: · The human variation UF of 50 is inconsistent with EPA guidance. · The UF of 3 for subchronic to chronic exposure is unneeded because the dosing period in the animal study was chronic. · The LOAEL to NOAEL UF is unneeded because two of the three studies relied upon provided NOAELs or their equivalent. · The UF for animal to human extrapolation is unneeded because there is strong evidence that humans are less sensitive to the effects of TCE

	than are the most sensitive rodent test species. • Using a modifying factor to account for background exposure in the course of deriving an RfD is clearly inappropriate in risk assessment (such factor might be applied, if determined to be needed, in the risk management phase).
NTP's proposed listing of naphthalene as a carcinogen	NTP's proposal to list naphthalene as "reasonably anticipated to cause cancer in humans" is contrary to NTP's own guidelines in that: • There is no "limited evidence of carcinogenicity from studies in humans" because the only suggestions of an association between naphthalene exposure and cancer are seriously confounded. • There is not "sufficient evidence of carcinogenicity from studies in experimental animals" because the evidence is in one species, not multiple species; the evidence is at one tissue site, not multiple tissue sites; the evidence is from one route of exposure, not multiple routes; the evidence does not show an unusually high tumor incidence or other unusual characteristic; and there are "compelling data indicating that [naphthalene] acts through mechanisms which do not operate in humans."
EPA's new RfC for naphthalene	•EPA derived the RfC by applying an uncertainty factor of 3000 to a Human Equivalent Concentration of 9.3 mg/m^3, a value that is over 5 times lower than the occupational standard (TWA-TLV = 50 mg/m^3). •The naphthalene RfC overstates toxicity because of undue sensitivity of the animal model, the misapplication of uncertainty factors, and the stark contrast between the RfC and real-world exposure data (the RfC of 0.003 mg/m^3 is virtually identical to the background concentration for the chemical of 0.0052 mg/m^3)..
Hoboken, New Jersey, industrial building remediation	• EPA requiring remediation of building to unattainable standard of 0.44 µg mercury/m^3 of air even though OSHA standard is 100 $µg/m^3$, ACGIH standard is 25 100 $µg/m^3$, WHO standard is 25 100 ug/m^3, and lowest standard in any of 16 other countries is 20 $µg/m^3$. • EPA standard allegedly imposed to protect for workplace exposure, but calculated using residential assumptions. • EPA standard imposed even though the most reliable worker exposure studies show that 25 $µg/m^3$.is adequately protective.
PCB TSCA "Megarule"	• 1998 PCB remediation standards based on CSF of 4.0 $(mg/kg/day)^{-1}$ even though in 1996 EPA lowered the CSF to a maximum of 2.0 $(mg/kg/day)^{-1}$.

Gas Turbine Ass'n petition to "delist" gas turbines from MACT	⬩ Inconsistent with guidance issued by EPA and other agencies, EPA requiring risk assessment to be performed using assumption that cancer and noncancer risks of chemicals are additive without any showing that the chemicals effect same target organ through same mechanism.
Use of Toxicity Equivalency Approach	⬩ EPA advocating use of the "Toxicity Equivalency" ("TEQ") approach to assessing PCB toxicity even though approach is based on the unproven assumption of additive toxic effects, a significant amount of data indicate that TEFs are not additive, and TEQ approach substantially overpredicts the cancer and noncancer toxicity of PCBs.
Ecological Risk Assessment ("ERA")	
Toxicity Assessment	
EPA recommended "weight-of-evidence" approach to derive TRVs	Although recent EPA guidance discusses four methods that may be used to derive Toxicity Reference Values ("TRVs"), it recommends use of a method, the so-called "weight-of evidence" method, that will not generate TRVs that are appropriate for individual sites: ⬩ The method, which derives a TRV as the geometric mean of a variety of studies, may not use measurement endpoints that are relevant to key ecological receptors at particular sites (i.e., it generates generic TRVs). ⬩ For example, when evaluating studies for a weight-of-evidence assessment, less weight should be given to studies that evaluate the toxicity of the given chemical to receptors that are not found at the particular site or whose chemical form may not be relevant to the site-specific form.
EPA avian TRV for dioxin for the Hudson River	⬩ On three occasions, EPA reviewed the toxicological data relevant to deriving an avian TRV for dioxin and decided that the TRV should be approximately equal to the NOAEL for dioxin from Nosek et al. (1992) study, which involved a 10-week exposure period. In so deciding, EPA determined that no subchronic to chronic uncertainty factor was needed because the study involved exposure throughout a critical life stage (reproduction). ⬩ Despite this precedent, EPA, in its ecological assessment for the Hudson River, used a TRV that was approximately 10 times lower based on the unwarranted assumption that the 10-week exposure period represented subchronic exposure.
EPA otter TRV for PCBs for the Hudson	For Hudson River otters, EPA developed TRVs based on data for mink. The TRVs are not scientifically supportable because the study used to

River	develop the TRVs included confounding exposures to pesticides and the authors did not attempt to segregate the potential contribution of the pesticides to the evaluated endpoint (kit survival).
Risk Characterization	
Fox River ERA	· Although an ERA is to consist of a "screening level" assessment followed by a detailed "baseline" ERA using site-specific data, final Fox River ERA is little more than a screening level assessment. · Although the ERA cites some of the voluminous site-specific ecological data that have been compiled, it ignores those data in arriving at its final conclusions.
Hudson River ERA	The Hudson River Revised Baseline ERA should not have been used to determine remedial action because the approach employed was designed for screening-level applications. · On behalf of EPA, Eastern Research Group coordinated a review of the ERA by seven independent peer reviewers. · Peer review group sharply criticized EPA's work product, concluding that EPA's ERA represented a screening-level effort and providing EPA specific recommendations to reduce the conservatism of and improve the ERA. · Peer review group unanimously agreed that EPA's characterization of the ecological setting was inadequate: "[W]ithout a description of the habitats, the species occupying the Hudson River, and the spatial and temporal use of habitats by species considered in the conceptual site model, the reviewers did not think it was possible to defend the risk characterization. . . . · EPA either failed to implement these recommendations, implemented the recommendations incorrectly, or made offsetting changes to the recommendations that resulted in little reduction to the level of conservatism.
PCB Worm Tissue Criterion for the Historic Area Remediation Site	In October 2002, EPA developed a proposed PCB worm tissue criterion for the "Historic Area Remediation Site" ("HARS"). The criterion is to be used to determine the suitability of dredged material for use as remediation material. The 113 ppb criterion is based on a number of overly conservative assumptions, including: · 100% of fish consumed by New Jersey anglers are sport-caught saltwater finfish.

	・100% of the fish consumed are caught at the HARS. ・ All species consumed by recreational anglers are available at the HARS. ・Anglers fish consistently every year for 70 years. ・There is no loss of contaminants due to cooking methods. EPA ignored comment raising these issue, and promulgated the 113 ppb criterion in March 2003.
Tier II Great Lakes Initiative water quality criteria	In the Great Lakes Water Quality Initiative ("GLWQI"), EPA proposed a two-tiered approach to deriving water quality criteria. A Tier I water quality criterion is derived when specific data requirements are met . These data requirements are identical to those that EPA has used historically as the minimum requirements for calculation of ambient water quality criteria. Under the GLWQI regulations, a Tier II water quality value can be derived if the data required to derive a Tier I value are not available, or if the data are not of high quality. Because Tier II criteria are to be derived based on incomplete or inferior data, EPA builds in several levels of conservatism in the calculations. The approach can result in extremely low values, particularly when only a few acceptable toxicity studies are available, because the conservatism in the Tier II value increases as the number of suitable studies decreases. For example: ・ A comparison of chronic Tier I values for nine metals to their corresponding Tier II values show that the Tier II values overestimate the Tier I values from 3 to 16,000 times. ・Use of the Tier II approach to develop a criterion for sodium chloride resulted in a criterion lower than naturally occurring levels.

6. Improving EPA Risk Assessments to More Accurately Estimate Risk

The issues presented in these comments are not new. As early as 1983, the National Research Council identified serious problems with the manner in which federal agency risk assessments were being conducted and recommended improvements to the process (NRC, 1983). The NRC's "Recommendation 1" was that as follows:

> *Regulatory agencies should take steps to establish and maintain a clear conceptual distinction between assessment of risks and the*

> *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 goal of risk assessment is to describe, as accurately as possible, the possible health consequences of changes in human exposure to a hazardous substance; the need for accuracy implies that the best available scientific knowledge, supplemented as necessary by assumptions that are consistent with science, will be applied.*

NRC (1983).

These recommendations went largely unheeded. Twelve years ago, the Executive Office of the President ("EOP") reviewed federal agency risk assessments and concluded that risk assessments were unduly conservative and inappropriately imported policy judgments into what should be the science of risk assessment:

> *Unfortunately, risk-assessment practices continue to rely on conservative models and assumptions that effectively intermingle important policy judgments within the scientific assessment of risk. Policymakers must make decisions based on risk assessments in which scientific findings cannot be readily differentiated from embedded policy judgments. . . .*
>
> * * *
>
> *The continued reliance on conservative (worst-case) assumptions distorts risk assessment, yielding estimates that may overstate likely risks by several orders of magnitude. Many risk assessments are based on animal bioassays utilizing sensitive rodent species dosed at extremely high levels. Conservative statistical models are used to predict low-dose human health risks, based on the assumption that human biological response mimics that observed in laboratory animals. Worst-case assumptions concerning actual human exposure are commonly used instead of empirical data, further exaggerating predicted risk levels.*
>
> Conservative biases embedded in risk assessment impart a substantial "margin of safety". The choice of an appropriate margin of safety should remain the province of responsible risk-management officials, and should not be preempted through biased risk assessments. Estimates of risk often fail to acknowledge the presence of considerable uncertainty, nor do they present the extent to which conservative assumptions overstate likely risks. Analyses

> of risk-management alternatives routinely ignore these uncertainties and treat the resulting upper-bound estimates as reliable guides to the likely consequences of regulatory action. Decision makers and the general public often incorrectly infer a level of scientific precision and accuracy in the risk-assessment process that does not exist.
>
> Conservatism in risk assessment distorts the regulatory practices of the Federal Government, directing societal resources to reduce what are often trivial carcinogenic risks while failing to address more substantial threats to life and health.

EOP (1991).

These three problems – policy judgments being inserted into the risk assessment phase, unjustified, overly conservative assumptions, and failure to acknowledge uncertainty – hamper risk assessment to this day, and have real consequences for human health, the environment, the federal budget, and private sector resources. Poorly done risk assessments can direct attention away from the actual source of risk. They also can lead to unnecessary expenditures of significant resources on insignificant risks, thereby reducing the resources available to address the significant risks. As new toxicogenomic tools are used, biased risk practices could damage the credibility of such tools, which otherwise might offer a solution to a number of the current risk assessment quandaries.

Fortunately, as discussed below, the means to improve EPA risk assessment are widely known, and implicit in EOP's 1991 critique of federal agency risk assessment. What continues to be lacking is the implementation of state-of-the-science knowledge and a discontinuation of widely discredited practices of the past. Leadership from OMB is needed to show other federal agencies, and particularly EPA, that better risk assessment and risk management decisions are possible and desirable. The Council's prescription for improving EPA risk assessment involves three fundamental changes in how risk assessment is performed today and several smaller changes that will assist in implementing the fundamental changes. These three key changes, as articulated in the 1991 EOP critique, are:

⇒ *EPA risk assessments must not "intermingle important policy judgments within the scientific assessment of risk*." Rather, the "choice of an appropriate margin of safety should remain the province of responsible risk-management

officials, and should not be preempted through biased risk assessments." This principle is simple – risk assessments should aspire to the greatest extent possible to be objective scientific exercises that seek to realistically estimate risk. Risk management comes later, and must be fully and transparently distinguished from risk assessment if the practice of risk assessment is to have scientific credibility.

⇒ *Risk assessments should not continue an unwarranted reliance on "conservative (worst-case) assumptions" that distort the outcomes of the risk assessment, "yielding estimates that may overstate likely risks by several orders of magnitude."* Further, worst-case assumptions concerning actual human exposure should not be used "instead of empirical data," because they further exaggerate predicted risk levels. In short, risk assessments should use real data to the extent feasible; be as accurate, precise and realistic as possible; and should not seek to embed conservative policy preferences into what should be a policy-neutral estimation of risk. In cases where realistic knowledge/data concerning a risk scenario is unavailable, assumptions are required, but to the extent that conservative assumptions are used in a risk assessment they must be clearly articulated for risk managers so that they fully understand how the analysis was performed, and where it may be overestimating risks.

⇒ *Risk assessments should "acknowledge the presence of considerable uncertainty" and present the extent to which conservative assumptions may overstate likely risks.* They should not "routinely ignore these uncertainties and treat the resulting upper-bound estimates as reliable guides to the likely consequences of regulatory action." Risk assessments should directly assess the impact of each choice or assumption and clearly communicate how these choices impact likely risks.

The three fundamental changes described above should be applied to all phases of risk assessment, including toxicity assessment and exposure assessment. The following hypotheticals

should assist in elucidating exactly how toxicity and exposure assessments would incorporate these fundamental changes and how these tools can be applied. The point of these hypotheticals, which admittedly are somewhat simplistic, is to illustrate the difference between risk assessment and risk management, and to make clear the dangers of compiling layers of conservative assumptions in the course of risk assessment rather than leaving risk management decisions to the risk manager.

Exposure Assessment.

Assume the following hypothetical:

> *The Brown landfill is a lined landfill with leachate collection that was closed, capped and fenced in 1998. The existence of groundwater contamination (primarily elevated concentrations of Chemical Y) led to a Remedial Investigation/Feasibility Study that suggested a need for groundwater remediation. Remediation is now under way pursuant to a Consent Order with the landfill operator. Pursuant to the Consent Order, groundwater will be pumped from the subsurface, treated along with landfill leachate, and discharged pursuant to an NPDES permit. The cap will be inspected and, if needed, repaired, on a semi-annual basis. The cleanup has been funded for a period of 30 years. The landfill is in an industrial area and there is no reason to believe that the property will ever be used for any other purpose.*
>
> *Soil surrounding the landfill is contaminated with high concentrations of Metal Y, which is not volatile. Modeling shows that leachate from the soil will be collected by the groundwater wells and/or leachate collection system.*

Given these assumed facts, and applying the above principles, what exposure factors should be used in assessing the risk from the presence of Chemical Y in soils at the Brown landfill?

The Council submits that in this hypothetical there is only one plausible exposure pathway – soil ingestion by landfill workers. Dermal exposure is not a potential pathway because HAZWOPER requires that workers in a landfill environment wear protective clothing. The exposure frequency should probably be no more than twice a year – apparently the only time the landfill property needs to be entered. The exposure duration should be based on the estimated number of years a landfill worker might reasonably be employed in that capacity. The soil ingestion rate should be

no greater than EPA's default value of 50 mg/day. There is no reason to posit a construction scenario; thus a higher ingestion rate is unnecessary.

An EPA risk assessor would probably not agree with the above exposure assessment. Based on the Council's experience, an EPA risk assessor would make various assumptions regarding future occurrences that might increase human exposure to Chemical X at the Brown landfill, assume what the exposures might be under those assumptions, and calculate risk estimates based on assumptions regarding future conditions. For example, the risk assessor might assume that the fence would be torn down and trespassers would visit the landfill twice a week (even though the site will be managed for the next 30 years). The assessor might even assume that someday, when the landfill is finished subsiding and venting methane, if will become a residential neighborhood, and children will play on the contaminated soil 5 days per week.

Adopting such implausible and unscientific assumptions is not a proper function in risk assessment. The likelihood that the fence will be torn down or the landfill turned into a residential neighborhood cannot be calculated – they are not scientific issues. Rather, they are predictions about the future that raise policy issues such as how much money should be expended to protect against the low chance that a hypothetical trespasser 30 years in the future will like to visit the Brown landfill. Or about how much money should be spent to protect against the unlikely possibilities that someone, someday, will want to develop a residential community on a landfill and that, at that time, the government will allow it. These are decisions that are not within the purview of a risk assessment professional. They are policy issues that need to be addressed – openly and in detail – at a high level within EPA with, at minimum, the concurrence of the Regional Administrator.

Toxicity Assessment.

Assume the following hypothetical:

> *In White (1997)[11], rabbits were subjected to lifetime exposure to airborne concentrations of 0.1, 1, and 10 mg/m^3 of Chemical X. All the rabbits died at the 10 mg/m^3 exposure level within two*

[11] Note that the White, Jones and Brown studies referred to in the hypotheticals are themselves hypothetical.

days. None of the rabbits died at the 1 mg/m^3 exposure level, although there was some evidence of subtle neurological effects. No effects were seen at the 0.1 mg/m^3 exposure level.

Widget manufacturing workers were exposed to high airborne concentrations of Chemical X in the 1940s, but exposure levels were not documented. An epidemiological study of some of these workers at a plant in Chile (Smith, 2000) found a statistically significant excess of age-related neurological deficits.

In the United States, worker exposure to Chemical X was likely never as high as in Chile because U.S. plants used closed top widget immersion baths. Use of Chemical X was phased out between 1950 and 1965 as a dry process was developed for widget production. Occupational exposure measurements of workers at one U.S. plant in 1964 revealed that they had an average Chemical X exposure at that time of 3 mg/m^3. An epidemiological study of these workers (Jones, 1999) found no statistically significant association between Chemical X exposure and any disease.

Given these assumed facts, and applying the above principles, what should the RfC be for Chemical X?

The Council submits that the following analysis should be applied. Clearly 0.1 mg/m^3 is a NOAEL and 1 mg/m^3 is a LOAEL for Chemical X in rabbits. Jones (1999) suggests 3 mg/m^3 is a NOAEL for Chemical X in humans. Smith (2000) suggests that Chemical X may cause serious adverse effects in humans at levels that are probably significantly higher than 3 mg/m^3. Given the uncertainty that 3 mg/m^3 is truly the NOAEL for Chemical X in human since is based on only one epidemiological study, the point of departure for deriving the RfC for Chemical X should be the rabbit LOAEL 1 mg/m^3.

Given the above facts, the Council submits that the uncertainty factor analysis set forth in Table 3 should be applied:

TABLE 2 – THE COUNCIL'S RECOMMENDED UNCERTAINTY FACTOR ANALYSIS FOR DERIVING AN RfC FOR CHEMICAL X		
Source of Uncertainty	**Factor**	**Rationale**
Use of LOAEL rather than NOAEL	None	Although a LOAEL was used as the starting point, Smith (2000) suggests that the human NOAEL is higher than 1 mg/m^3. Therefore, use of 1 mg/m^3 as the starting point has already applied, in effect, a UF of 3.
Interspecies extrapolation	None	Not needed because the evidence is that rabbits are more sensitive than people to the effects of Chemical X.
Variation in human sensitivity	3	Generally, differences in sensitivity among human subpopulations are not large; usually the difference is less than a factor of 2 to 3, but occasionally is exceeds 5 (Price et al., 1999). Where the critical effect of the compound is well understood in humans and the sensitive population has been identified, the uncertainty factor for human variation can be reduced to as low as 1, as in the current RfDs for nitrate and fluorine (soluble fluoride) (Cicmanec and Pourier, 1995; EPA, 2003).
Extrapolation from subchronic to chronic exposure	None	White (1997) was a chronic study.
Database uncertainty	None	The weight of the evidence from three studies suggests strongly that the level at which Chemical X causes adverse effects in humans is well above 0.33 mg/m^3 (rabbit LOAEL of 1 mg/m^3 divided by UF of 3 for variation in human sensitivity).
Total uncertainty factor	3	

The RfC is therefore 0.33 mg/m^3 (rabbit LOAEL of 1 mg/m^3 divided by UF of 3).

Note that an EPA risk assessor would almost certainly establish a more stringent RfD based on this set of facts. This is true because EPA risk assessments universally violate the fundamental principles set forth above – they seek to be highly conservative, they apply policy judgments,

and they fail to appreciate the extent to which they compound conservatism. They do not seek to predict risk as accurately as possible.

For purposes of comparison, the Council sets forth below how we believe an EPA risk assessor would use the above hypothetical facts to derive an RfC for Chemical X:

EPA would agree that 0.1 mg/m^3 is a NOAEL and 1 mg/m^3 is a LOAEL for Chemical X in rabbits. EPA would also agree that Jones (1999) suggests 3 mg/m^3 is a NOAEL for Chemical X in humans. EPA would agree that Smith (2000) suggests that Chemical X may cause serious adverse effects in humans at levels higher than 3 mg/m^3, although how much higher is unknown. EPA would feel that there is substantial uncertainty regarding whether 3 mg/m^3 is truly the NOAEL for Chemical X in humans since it is based on only one epidemiological study. Thus, EPA would decide that the point of departure for deriving the RfC for Chemical X should be the rabbit NOAEL 0.1 mg/m^3.

Given the above facts, the Council believes that EPA would apply the uncertainty factor analysis set forth in Table 3:

TABLE 3 – THE COUNCIL'S PROJECTION OF EPA'S UNCERTAINTY FACTOR ANALYSIS FOR DERIVING AN RfC FOR CHEMICAL X		
Source of Uncertainty	**Factor**	**Rationale**
Use of LOAEL rather than NOAEL	None	Not needed because NOAEL used as starting point.
Interspecies extrapolation	3	Insufficient data to assume that rabbits are more sensitive than humans. (Based on experience, EPA might very well use a factor of 10.)
Variation in human sensitivity	10	Since no data are available for Chemical X, full factor of 10 applied.
Extrapolation from subchronic to chronic exposure	None	White (1997) was a chronic study.
Datebase uncertainty	10	Relatively few studies exist, so full factor of 10 is appropriate.
Total uncertainty factor	300	EPA virtually always multiplies, rather than adds, uncertainty factors.

EPA would therefore derive an RfC for Chemical X of 0.00033 mg/m^3 (rabbit NOAEL of 0.1 mg/m^3 divided by UF of 300).

Table 4 presents a comparison of the two hypothetical derivations of RfCs for Chemical X:

TABLE 4 – COMPARISON OF COUNCIL's AND EPA's RfCs FOR CHEMICAL X			
Factor	**Council**	**EPA**	**Reason for Difference**
Toxicity Assessment			
Point of Departure	rabbit LOAEL of 1 mg/m^3	rabbit NOAEL of .1 mg/m^3	Difference in weight given to negative epidemiological study
Exposure Assessment			
Use of LOAEL rather than NOAEL	None	None	None, although EPA would have applied a factor of 3 – 10 if it had

			used LOAEL as point of departure
Interspecies extrapolation	None	3	Difference in weight given evidence showing that rabbits are more sensitive than humans
Variation in human sensitivity	3	10	EPA almost always uses a factor of 10; the Council advocates 3 because the critical effect of Chemical X is well understood in humans, albeit in a population of workers who may have been healthier than other members of the general population. Because there clearly were no adverse effects in this human study, an uncertainty factor of 3 should be more than adequate to account for variation in human sensitivity.
Extrapolation from subchronic to chronic exposure	None	None	None
Datebase uncertainty	None	10	Council believes database to be unambiguous – human effects appear well above the Council RfD value. EPA applies full factor because few studies exist.
Total uncertainty factor	3	300	
Toxicity Characterization			
RfC	0.33 mg/m^3	0.00033 mg/m^3	

* * *

Although careful attention to and application of the three principles discussed and illustrated above should enable a risk assessor to produce an unbiased risk assessment, several tools exist that can be used to assist in this effort. All of these tools should assist the risk assessor to avoid including policy judgments within the risk assessment (leaving that to the risk manager) and

producing a risk assessment that, to the extent possible, neither overstates nor understates risk. These tools are summarized below. EPA guidance supporting the use of these tools is footnoted.

- *Risk assessors should present managers with a range of risk scenarios and fully disclose the plausibility of each to facilitate the risk manager's informed policy choices*. OMB must direct agencies in their risk assessments to consider multiple scenarios and to fully account for the plausibility or likelihood of each. Within this process, agencies must consider the highly unlikely but plausible worst-case, the expected value or mean estimate of risk, and the reasonable best-case outcomes, without unduly emphasizing worst-case hypothetical scenarios. In presenting risk managers with a range of possible outcomes, accurately weighted for their likelihood, the goal of risk assessors should be to support the managers in making fully informed choices about both the appropriate degree of conservatism or precaution to adopt and the extent to which such choices may entail tradeoffs among other important factors (i.e., to facilitate the risk manager's informed consideration of benefits and costs). In comprehensively disclosing the features of their assessments, risk assessors must provide the empirical basis or scientific rationale for any assumption, conservative or precautionary policy choices used in a given scenario. They must also fully explain the implications of choosing a particular policy, including the countervailing risks and other effects that might arise directly or indirectly from a decision based on such policy choices. While default assumptions are required to fill data gaps and address uncertainties that arise in the conduct of a risk assessment, it is the risk manager's responsibility to ultimately decide how to address limitations in the risk assessment through additional safety factors and other policy decisions. Risk assessments must serve, not usurp, this process.

  Greater reliance on certain tools can facilitate the risk manager's role in making choices. For example, although EPA has nominally endorsed Monte Carlo and other stochastic methods in conducting risk assessments [EPA 1997], the Agency has not defined the process or data would make these tools truly effective. EPA is in a unique position to evaluate, on a scientific basis, the quality of data for use in this application, but the Agency has not addressed this topic. As an example, Stanek and

Calabrese developed a soil ingestion distribution for children in a study that was funded by the EPA, but EPA has yet to endorse this data set as appropriate for this application. [Stanek, 2000, 2001] This reticence on the part of the Agency sends a signal that is contrary to EPA's 1997 policy, when in reality this is an area where EPA can clearly drive the science forward rather than continuing to use methods that are outdated and lead to the mis-allocation of limited resources to correct problems that will have little or no real impact on improving the public health. Probabilistic methods can also be applied to toxicity data. When data are clearly available, as in the case of higher-tier risk assessments, a probabilistic approach is the most scientifically appropriate. This will provide the risk manager with the proper frame of reference for making decisions, as opposed to the policy laden deterministic approach currently in place.

- *Agencies should assess scientific evidence using a weight-of-the-evidence process that is consistent, comprehensive, balanced, and reproducible*. Although EPA describes the approach it uses in its toxicity assessments (and sometimes in performing exposure assessments) as a weight-of-the-evidence process, in fact the Agency does not follow consistent, comprehensive, balanced, and reproducible procedures that external parties can clearly follow and understand. Such procedures assist the risk assessor in deciding which data, both positive and negative, should be given more weight, and in determining how disparate data can be combined to reach a rational and scientifically supportable conclusion. To be useful and understandable to external parties, the process EPA employs must be a more formal and transparent weight-of-the-evidence process, such as the approach developed by Klimisch, et al., for evaluating data quality (Klimisch 1997), the Bradford-Hill causation criteria cited below, and other such approaches that can make it clearer how EPA risk assessors judged the evidence they considered. A formal process would assign weights to data or apply carefully defined evaluation criteria to assist the risk assessor in deciding which data should be given more weight and in determining how disparate data can be combined to reach a rational and scientifically supportable conclusion. In addition, EPA's weight-of-the-evidence process must:

⇒ Place greater emphasis on human studies. Although EPA often states that human studies (including epidemiological studies) should be given more weight than animal studies, in practice the Agency does not consistently follow this policy. In particular, EPA sometimes dismisses epidemiological studies of any quality that do not show positive associations and accepts with little resistance studies that yield positive associations irrespective of their scientific quality. Epidemiological studies of highly exposed occupational cohorts provide important information on the human toxicity of chemicals and should inform EPA toxicity assessments to a much larger extent than at present.

⇒ Use causation analysis. Causation analysis, sometimes referred to as application of the Hill Criteria (Bradford-Hill 1966), should be used to evaluate whether exposure to a particular chemical may cause an increased risk of disease. Specifically, causation analysis should be applied to a group of studies that have investigated potential associations between exposure to a particular chemical and a specific disease endpoint.

The Council notes that EPA's 1996 Proposed Guidelines for Carcinogen Risk Assessment unequivocally endorse the weight-of-evidence approach for evaluating epidemiological data applicable to a particular chemical and describes "well-accepted criteria for causation" that should be used in such an approach. At least 10 cause and effect analytical criteria have been proposed, according to EPA's document, though only six are described as "fundamental" criteria (EPA 1996, pD-9). EPA's 1996 proposal reads:

> *Analyzing the contribution of evidence from a body of human data requires examining available studies and weighing them in the context of well-accepted criteria for causation. A judgment is made about how closely they satisfy these criteria, individually and jointly, and how far they deviate from them. Existence of temporal relationships, consistent results in independent studies, strong association, reliable exposure data, presence of dose-related responses, freedom from biases and confounding factors, and high level of statistical significance are among the factors leading to increased confidence in a conclusion of causality. Generally, the weight of human evidence increases with the number of adequate*

> *studies that show comparable results on populations exposed to the same agent under different conditions. The analysis takes into account all studies of high quality, whether showing positive associations or null results, or even protective effects. In weighing positive studies against null studies, possible reasons for inconsistent results should be sought, and results of studies that are judged to be of high quality are given more weight than those from studies judged to be methodologically less sound. Generally, no single factor is determinative. For example, the strength of association is one of the causal criteria. A strong association (i.e., a large relatively [sic] risk) is more likely to indicate causality than a weak association. However, finding of a large excess risk in a single study must be balanced against the lack of consistency as reflected by null results from other equally well designed and well conducted studies. In this situation, the positive association of a single study may either suggest the presence of chance bias or confounding, or reflect different exposure conditions. On the other hand, evidence of weak but consistent associations across several studies suggests either causality or the same confounder may be operating in all of these studies."(EPA 1996, pD-6-7)]*

- *Agencies should accept site- or chemical-specific data*. Although EPA recommends use of site-or chemical-specific data, it often does not accept their use, requiring instead that conclusive or unambiguous evidence be provided before a default value can be superseded. OMB should direct agencies to use site- or chemical-specific information first, and if these data are unavailable, an agency may consider a safety or default value consistent with the above recommendation.

Where possible, risk assessments for specific sites should be based on reasonable and realistic exposure measurements or estimates for the site in question, not default or assumed values.[12] In determining the appropriate exposure area to be evaluated, the entire area that is equally likely to be contacted by the receptor should be considered, not just the contaminated portion of that area; otherwise, exposures will be

[12] See EPA (1992), at 93:

> The Exposure Factors Handbook is being updated to encompass additional factors and to include new research data on the factors currently covered. It also provides default parameter values that can be used when site-specific data are not available. Obviously, general default values should not be used in place of known, valid data that are more relevant to the assessment being done.

overestimated.[13] In determining exposure point concentrations for that area, agencies should estimate representative average concentrations using an appropriate statistical technique, rather than using maximum concentrations or statistical techniques that overestimate the true average.[14] In deriving values for other exposure parameters (e.g., exposure frequencies, amount of skin surface area exposed, food consumption rates, etc.), estimates and assumptions should be reasonable and realistic for the site, taking into account current and reasonably foreseeable site uses and conditions.[15] Use of probabilistic techniques to estimate exposures should be encouraged where appropriate.[16]

---

[13] See EPA (1989), at 6-26:

> When evaluating chemical contamination at a site, it is important to review the spatial distribution of the data and evaluate it in ways that have the most relevance to the pathway being assessed. In short, consider where the contamination is with respect to known or anticipated population activity patterns. Maps of both concentration distribution and activity patterns will be useful for the exposure assessment. It is the intersection of activity patterns and contamination that defines an exposure area. Data from random sampling or from systematic grid pattern sampling may be more representative of a given exposure pathway than data collected only from hot spots.

[14] EPA (2001) defines the reasonable maximum exposure ("RME") as "[t]he highest exposure that is reasonably expected to occur at a site." EPA guidance (1992, 2001) recommends that, in developing the RME or "high end" exposure scenario, a combination of average and high-end exposure assumptions be combined to develop a "reasonable" estimate of maximum exposure rather than the worst-case hypothetical exposure imaginable.

[15] See EPA (1995) ("Land use assumptions affect the exposure pathways that are evaluated in the baseline risk assessment. Current land use is critical in determining whether there is a current risk associated with a Superfund site, and future land use is important in estimating potential future threats.").

[16] See EPA (2001). The guidance describes a tiered process for conducting HHRAs at complex sites. Beginning with deterministic point estimates, the analysis progresses, if warranted, to increasingly complex tiers of probabilistic analysis. The guidance makes clear that it is applicable to both human health and ecological risk assessments. Page xv states that "PRPs may submit work plans for probabilistic risk analyses for review during the risk assessment process or as required under legal agreements." The guidance also states that Monte Carlo analysis adds value whenever screening risk estimates exceed levels of concern and when the costs of remediation are likely to be high. Implicit in EPA's (2001) recommended progression from deterministic to probabilistic analysis is the realization that the point estimates are not an adequate means for basing risk management decisions at complex sites. According to EPA (2001) (at p.1-16):

> *The point estimate approach to risk assessment does not determine where the CTE or RME estimates lie within the risk distribution. … Without knowing what percentile is represented by the RME estimate, the risk manager might be unsure about the likelihood of the RME occurring or being exceeded in the receptor population and about what level of remedial action is justified or necessary. . . .*

As a specific application of this principle, realistic default factors should be used in EPA's Soil Screening Guidance instead of the overly conservative values currently in place. This is especially important in the case of the factors for the migration-to-groundwater exposure route, which end up suggesting that very small amounts of soil contamination could contaminate groundwater, and therefore require full-scale risk assessments for chemicals that should be screened out early in the risk assessment process.

EPA should approve, in practice, the use of site-specific values in risk assessment and thereby follow and conform to the intent of its own guidance. If EPA were to do this, then many high-quality studies would be performed, e.g. field studies, land use studies, recreational surveys, etc., so as to improve the quality and accuracy of human health and ecological risk assessments.

- *Agencies should fully implement the Information Quality Guidelines*. OMB should insist that federal agencies fully apply their Information Quality Guidelines in the course of conducting risk assessments, and should do so in a manner that is consistent with OMB's government-wide standards. Agencies should defer to studies that meet these guidelines and must set aside potentially influential information that is not transparent enough to be reproducible or data deemed of questionable utility or integrity. In addition, information quality and applicability must be the primary drivers for weight-of-the-evidence procedures, causation analysis, and the use of site- or chemical specific data (see items above).

Lastly, when appropriate – for instance, when a pesticide is engineered for deliberate toxicity or a potent chemical's widespread use may create broad exposure – OMB should recommend ecological risk assessments be conducted on effects at the population and/or community level rather than on individual receptors.

By characterizing variability with one or more input distributions, the output from the Monte Carlo simulation is a distribution of risks that could occur in that population. The central tendency of the risk distribution (e.g. arithmetic mean, geometric mean, 50th percentile) may be characterized as the CTE risk estimate. Similarly, the high-end of the risk distribution (e.g., 90th to 99.9th percentile) is representative of exposures to the RME individual.

EPA (1997) ecological risk assessment guidance recommends that potential ecological risks should be assessed at the population-level for all but threatened and endangered species. Although no explicit guidance is provided, this is typically accomplished through the use of measurement endpoints that are related to population effects (e.g., using Toxicity Reference Values based on growth or reproductive effects). However, in many EPA ecological risk assessments, the agency has defaulted to assessing effects on individual animals. This metric has no significance scientifically and is entirely useless as a basis for making risk management decisions.[17]

---

[17] See, e.g.:

EPA (1999), at 3,5:

> *Ecological risk assessments incorporate a wide range of tests and studies to either directly estimate community effects (e.g., benthic species diversity) or indirectly predict local population-level effects (e.g., toxicity tests on individual species), both of which can contribute to estimating ecological risk. Superfund remedial actions generally should not be designed to protect organisms on an individual basis (the exception being designated protected status resources, such as listed or candidate threatened and endangered species or treaty-protected species that could be exposed to site releases), but to protect local populations and communities of biota. Levels that are expected to protect local populations and communities can be estimated by extrapolating from effects on individuals and groups of individuals using a lines-of-evidence approach.*
>
> *Superfund risk assessments should use site-specific assessment endpoints that address chemical specific potential adverse effects to local populations and communities of plants and animals (e.g., reductions in populations of fish-eating birds, or reductions in survival, reproduction or species diversity of indigenous benthic communities).*

USEPA (1994), at 1:

> *The ecological risk assessment of a Superfund site nearly always requires some type of field study. At a minimum, some field study is necessary in order to identify organisms and habitats that may be at risk" and "Rather than studying individual organisms, field studies generally focus on populations or communities. Populations are groups of organisms belonging to the same species and inhabiting a contiguous area. Communities consist of populations of different species living together.*

USEPA (2001), at 9, 52. This document reviews many of the principles to develop ecological management objectives. It states than an "ecological risk assessment examines many different species and multiple levels of biological organization, from individual to population, community, and ecosystem. " It emphasizes the use of Case Studies to explain the management objectives process, developing discussions concerning the importance of "reducing the level of toxic substances and by protecting human health, restoring vital habitats, and restoring and

C. Balancing Precaution and Other Societal Interests

OMB's F.R. notice also requests comment on how the U.S. balances precautionary approaches to health, safety and environmental risks with other interests such as economic growth and technological innovation. In asking for comment on this issue, OMB raises a question with significance far beyond that of calculating benefits and costs of regulations for this nation. Technological innovation and economic growth are inseparable. When, despite reasonable measures to manage foreseeable risks, some government officials promote regulatory regimes that are so precautionary they thwart technological innovation, both the direct benefits of that innovation and the economic growth associated with it may be lost, or forestalled. This issue is discussed in a recent paper, "The True Cost of Precautionary Chemicals Regulation," which critiques the European White Paper "Strategy for a Future Chemicals Policy" and describes research that suggests the proposed strategy is "unrealistic and even unrealizable." (Durodié 2003)

For an economy – and an age – that is experiencing continual breakthroughs in biological sciences, nanotechnology, chemistry, and many other areas, it is imperative that our risk management systems not become an unreasonable obstruction to life-enhancing innovations. The key is to make reasonable judgments as to the appropriate balance between precaution and progress and to transparently and clearly describe the public interests being served by the decisions made, as well as the basis for the decisions. As has been said by local officials and

---

mainlining stable, diverse and self-sustaining populations."

USEPA (2002), at 10. This draft SAB document provides guidance on assessing ecological conditions in an environmental assessment context, but many of its components are relevant to ecological risk assessments that assess chemical exposures. For example, it defines species or population measures as:

> *Measures of the condition or viability of populations of species in an area are important indicators, yet monitoring the status of all species is impossible from a practical standpoint. To address this problem, a higher taxonomic level can be used, or a subset of species called focal species can be monitored. Focal species are selected because they exert a disproportionately important influence on ecosystem condition or provide information about the ability of the system to support other species. In addition, some species such as endangered, rare, sensitive, and game species) require attention because they are of direct interest to society.*

others, federal laws and regulations to advance one goal – for instance, environmental protection – should not be implemented without regard for other equally pressing social demands, such as police, fire, and other services.

Evolving tools for evaluating comparative risks and countervailing risks make a valuable contribution to a more holistic understanding of the tradeoffs inherent in many decisions and should be resorted to more routinely by federal agencies in setting priorities. Just as decision makers now consider environmental consequences when weighing the merits of many proposed activities, including, for instance, new housing and product development, so too should decision makers consider the impacts on economic and scientific progress when precaution is invoked as a reason to slow or halt an activity or technology.

A strong, effective, but efficient and fair risk management system is vital in a society increasingly reliant on numerous technologies, none of which are risk free but few of which we have needed to manage by resorting to bans. The Council strongly supports OMB's efforts to comprehensively evaluate the strengths and weaknesses of the current environmental, health and safety management system and to create a system that appropriately balances precaution in managing risks with other fundamental public interests.

* * *

IV. CONCLUSION

The Council is pleased to have had the opportunity to comment on the Office of Management and Budget's Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations. The Council appreciates OMB's interest and efforts in fostering accurate, balanced and cost-effective risk assessment and risk management. Although the Council fully understands that an appropriate degree of caution should be used in making risk management decisions, such caution should be applied transparently in the risk management phase, and not opaquely in the risk assessment phase. Risk assessments, for their part, must seek to provide the best and most objective estimates of risk possible, leaving policy judgments to the risk managers. If these fundamental changes are actually implemented by EPA, all sectors will benefit.

The Council hopes these comments will assist OMB and others in the Executive Office who are seeking to improve federal regulatory analysis and management. The Council looks forward to working with OMB on this and similar matters in the future.

V. REFERENCES

Bradford-Hill, A. 1996. The Environment and Disease: Association or Causation? Proc. Royal Soc. Med. 58:295

Burmaster, D.E. and J.H. Lehr, "It's Time to Make Risk Assessment a Science," Ground Water Monitoring Review XI(3):5-15 (summer 1991).

Cicmanec, J.L. and K.A. Poirier. 1995. Selected applications of reduced uncertainty factors in noncancer risk assessment. Hum. Ecol. Risk Assess. 1(5):637-640.

Durodié, Bill 2003. The True Cost of Precautionary Chemicals Regulation. Risk Analysis, Vol. 23, No. 2: 389-398.

EOP. Executive Office of the President of the United States. 1991. Regulatory Program of the United States Government. Current Regulatory Issues in Risk Assessment and Risk Management. (April 1, 1990 – March 31, 1991).

EPA. 1986. Guidelines for Carcinogen Risk Assessment. Federal Register 51(185):33992-34003.

EPA. 1989. Risk Assessment Guidance for Superfund, Volume I, Human Health Evaluation Manual (Part A). U.S. Environmental Protection Agency, EPA/540/1-89/002. December.

EPA. 1992. Guidelines for Exposure Assessment. Federal Register 57(104): 22888-22938.

EPA. 1994. ECO Update. Field Studies for Ecological Risk Assessment. Intermittent Bulletin, Volume 2, Number 3. Office of Solid Waste and Emergency Response. Publication 9345.0-05I, March.

EPA. 1995. Land Use in the CERCLA Remedy Selection Process. OSWER Directive No. 9355.7-04. May 25.

EPA. 1996. Proposed Guidelines for Carcinogen Risk Assessment. EPA/600/P-92/003C.

EPA 1997, Policy for Use of Probabilistic Analysis in Risk Assessments at the U.S. Environmental Protection Agency, Office of Research and Development, May 15, 1997.

EPA. 1999. Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites. Office of Solid Waste and Emergency Response. OSWER Directive 9285.7-28 P. October 7.

EPA. 2000. Presenter's Manual for: "Superfund Risk Assessment and How You Can Help." EPA/540/R-99/013. Office of Solid Waste and Emergency Response. July 2000.

EPA. 2001. Risk Assessment Guidance for Superfund: Volume 3 - (Part A, Process for Conducting Probabilistic Risk Assessment). Final. United States Environmental Protection Agency, Solid Waste and Emergency Response. EPA 540-R-02-002. December.

EPA. 2002. An SAB Report: A Framework for Assessing and Reporting on Ecological Condition. Science Advisory Board. EPA-SAB-EPEC-02-xxx. March. Draft.

EPA. 2003. Integrated Risk Information System (IRIS). National Center for Environmental Assessment.

Evans, John S., et al. 1992. Summary of Workshop to Review an OMB Report on Regulatory Risk Assessment and Management. RISK: Health, Safety & Environment. 3.71.

Hawley, J.K. 1985. Assessment of health risk from exposure to contaminated soil. Risk Analysis 5(4):289-302.

Klimisch, H.J., et al., 1997. A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data, Regulatory Toxicology and Pharmacology, 25:1-5.

Nosek, J.A., S.R. Craven, J.R. Sullivan, S.S. Hurley, and R.E. Peterson. 1992. Toxicity and reproductive effects on 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant hens. Journal of Toxicology and Environmental Health. 35: 187-198.

NRC. National Research Council. 1983. Risk Assessment in the Federal Government: Managing the Process. National Academy Press.

Price, P.S., C.L. Curry, P.E. Goodrum, M.N. Gray, J.I. McCrodden, N.W. Harrington, H. Carlson-Lynch, and R.E. Keenan. 1996. Monte Carlo modeling of time-dependent exposures using a Microexposure Event approach. Risk Anal. 16(3):339-348.

Price, P.S., R.E. Keenan, and B. Schwab. 1999. Defining the interindividual (intraspecies) uncertainty factor. Hum. Eco. Risk Assess. 5(5):1023-1033.

Stanek, E.J. III and Calabrese, E.J. 2000. Daily Soil Ingestion Estimates for Children at a Superfund Site. Risk Analysis 20:627-635.

Stanek, E.J. III and Calabrese, E.J., and Zorn M. 2001. Soil Ingestion Distributions for Monte Carlo Risk Assessment in Children. Human and Ecological Risk Assessment. Vol. 7, No. 2, pp 357-368.

Viscusi, W.K. and James T. Hamilton. April 1999. Are Risk Regulators Rational?

Evidence from Hazardous Waste Cleanup Decisions (Working Paper 99-2). (Research support was provided under Cooperative agreement No. CR-823604-01 from the U.S. EPA Office of Policy, Planning, and Evaluation).

VI. APPENDICES

APPENDIX 1 W.Kip. Viscusi & James T. Hamilton, Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions (AEI-Brookings Institute Joint Center for Regulatory Studies, Working Paper 99-2 (April 1999)

APPENDIX 2 AMEC Earth & Environmental, Development of a Revised Chronic Reference Dose for Polychlorinated Biphenyls (Aroclor 1254) Based on Empirical Data

APPENDIX 3 AMEC Earth & Environmental, Use of a Toxic Equivalency Quotient Approach Based on 2,3,7,8-TCDD to Evaluate Potential Carcinogenic Risks of PCBs (Sept. 26, 2001)

APPENDIX 4 Table of conservative default assumptions from EPA risk methodologies

APPENDIX 5 Examples of Risk Assessments that Grossly Overstate Risk

Executive Summary

ACC strongly supports the proposed Bulletin. Federal agency risk assessment policies and practices are currently uneven and lag behind the state of the art in risk assessment. The Bulletin distills almost a quarter century of consensus recommendations regarding risk assessment theory and practice from the National Academy of Sciences and similar august bodies. Application of its uniform, minimum standards would greatly improve agency risk assessment practices and bring them up to date.

The Bulletin is amply supported by OMB's legal authorities. By the same token, it would not impair agencies' ability to implement their statutory mandates. Nor would it impose significant additional burdens on agencies to implement. Indeed, by improving the quality and thus the defensibility of agency risk assessments, it could actually save agency resources.

The Bulletin appropriately applies to documents that are components of risk assessments. Many of these, such as IRIS values, are extremely influential, and it would be inefficient and unwise to defer application of the Bulletin until these values were used in some full risk assessment.

OMB should state more clearly that compliance with relevant provisions of the Bulletin is mandatory on agencies. It should also make clear that the Bulletin applies to independently-conducted risk assessments when they are disseminated by federal agencies. The Bulletin should apply to risk assessments prepared by an agency and relied on it to make permitting and licensing decisions. The Bulletin should not allow for waivers, only deferrals. OMB should clarify that "regulatory analysis" includes any analysis used to support rulemaking. A more comprehensive inventory of "influential" risk assessments would aid OMB in implementing the Bulletin. OMB should better articulate how its minimum standards would apply to ecological risk assessments.

OMB should more clearly and strongly establish the linkage between the Bulletin and OMB's guidance under the Information Quality Act. In particular, OMB should expressly reference the Safe Drinking Water Act objectivity provisions in Section IV.4 of the Bulletin, probably its most important provision. The preamble should be revised to confirm that agencies "shall," not "should," follow the Peer Review Bulletin. All risk assessments should be reproducible. The Bulletin could more clearly demonstrate how good risk assessments are crucial to good benefit-cost analysis.

The Bulletin should establish a general expectation that agencies will seek public participation in risk assessment. Early and frequent public participation is the best way to ensure that risk assessments are based on the best available information and are appropriately scaled and oriented to the relevant questions.

OMB should establish procedures for implementation of the Bulletin, primarily by requiring agencies to publish implementation plans that explain how they will revise their existing risk assessment policies in light of the Bulletin's requirements.

OMB should require that agencies (i) update and replace default assumptions and risk assessments as relevant and scientifically plausible information becomes available, (ii) disseminate those new assumptions and assessments, and (iii) cease relying on or otherwise disseminating older assumptions and assessments after such updates.

The Bulletin's provisions regarding comparisons with risk assessments conducted by qualified scientific organizations on the same topic should promote beneficial competition among such organizations. In this connection, OMB should require agencies to discuss the extent to such other assessments meet the requirements of the Bulletin, and should clarify that only truly independently-conducted risk assessments require consideration.

Screening-level risk assessments should only be used to determine whether fuller assessment is warranted. Even screening-level risk assessments should present central tendency results and discuss uncertainty.

Additional guidance is needed on how to make risk assessments "commensurate" with risks.

The Council strongly supports the requirement that judgments regarding whether a specific effect is adverse be specifically identified and justified. The Council also recommends that non-adverse effects not become the basis for regulatory decision-making unless and until their utility and reliability have been demonstrated.

The Bulletin should contain stronger provisions regarding reliance on the best-available information, absence of bias, and transparency. General requirements for risk characterization should also be strengthened.

Discussion

I. The Bulletin Is Clearly Needed and Can Greatly Improve Federal Agency Risk Assessments Without Impairing Agency Missions

A. Risk Assessment in the Federal Government Would Benefit from Uniform Standards and Oversight

1. Risk is the cornerstone of much federal regulation

The concept of risk is fundamental to federal law and regulation affecting the environment; workplace health and safety; food, drugs and cosmetics; highway safety; aviation safety; and homeland security – among other topics. A very substantial portion of the U.S. Code and the Code of Federal Regulations is devoted to risk-based regulation. Given the extraordinary amount of private and public sector expenditures driven by these authorities, it is essential that federal agencies assess the relevant risks in a consistent,

efficient and accurate fashion. Agency resources for conducting risk assessments are also limited, and thus should be used efficiently.

2. Agency risk assessment practices vary in quality and currentness

Unfortunately, agency risk assessment practices continue to suffer from a range of features that have been identified – in many cases, years ago — as problematic. These features systematically exaggerate actual risks and thereby seriously compromise the value of risk assessments as inputs to regulations and regulatory impact analyses. ACC documented the problems associated with these practices in its 2003 submission to OMB noted earlier. In brief, however, ACC is most concerned about the following practices:

- *Intermingling of policy judgments with scientific assessments*. EPA freely acknowledges that it mixes risk management policy choices into the risk assessment process.[1]
- *Reliance on conservative worst-case assumptions, such as extreme and implausible estimates of exposure* (e.g., "maximally-exposed individual"). EPA frankly admits that it does this.[2]
- *Selective use of relevant test results.* A pattern of policy-biased selections is typically practiced, in which attention is focused narrowly on those results from toxicological or epidemiological studies that lead to the highest estimates of risk (or lowest estimates of an "acceptably safe dose"). Data sets are selected that display the effect(s) at the lowest dose level, for the most sensitive effect, in the most sensitive organ or tissue, all in the most vulnerable species, strain and gender. Conflicting evidence (specifically, evidence that tends to support a conclusion of lower or no risk) is customarily discounted or ignored.
- *Basing cancer risk estimates primarily on statistical upper-bounds* (of risk, at a specified dose) *or statistical lower-bounds* (of dose, for a specified level of risk).
- *Basing risk assessments on non-adverse effects* (i.e., adaptive changes, absent evidence for adverse consequences).
- *Failing to acknowledge the considerable uncertainty inherent in risk assessments and the degree to which that uncertainty is masked by use of assumptions*.
- *Requiring full-fledged risk assessments where screening assessments could generate sufficient information for the question at hand*.

Agencies often display strong resistance to updating risk assessments. For example, EPA was unwilling to revise its reference dose (RfD) for perchlorate ingestion – despite strong scientific data and a biologically plausible rationale which called into question EPA's RfD – until after (i) a comprehensive and independent workshop was held on the state of

[1] *See* EPA Office of the Science Advisor Staff Paper, *An Examination of EPA Risk Assessment Principles and Practices* (EPA/100/B-04/001) (Feb. 2004), § 2.1.3 ("These policy positions not only shape the risk assessment process, but are also a factor in the decision-making process outside of risk assessment.").
[2] *Id.* § 2.2.7 ("[EPA's Office of Air and Radiation] has not modified the assumption of 70-year, 24-hour per day, outdoor exposure[, even though] OAR recognizes that the majority of people do not reside outdoors and in one location for their entire lives.").

the science with respect to possible health risks from perchlorate exposure and (ii) an expert panel of the National Research Council (NRC) further reviewed the relevant and available scientific evidence on perchlorate (including the report from the state-of-the-science workshop). Following these events, EPA agreed to revise the Agency's proposed RfD for perchlorate.[3]

3. An OMB Bulletin on risk assessment could greatly improve the quality of agency risk assessments

Many of the problems just discussed could be reduced or eliminated by upgrading the quality, objectivity, utility and integrity of risk assessment practices across federal agencies. For example, many current practices are contrary to the Bulletin's prescription that risk assessments "neither minimize[e] nor exaggerat[e] the nature and magnitude of the risk." As another example, at the NRC conference considering the Bulletin, former OIRA Administrator John Graham noted that if EPA had initially performed a perchlorate risk assessment that conformed to the provisions of the Bulletin (for example, by providing central estimates of risk and applying conventionally-accepted criteria for specifying adverse effects), time-consuming and resource-intensive disputes among federal agencies could have been avoided, and a defensible RfD could have been issued more quickly and at lower cost to EPA, other federal agencies and interested parties.[4] The comments below further explain how many proposed Bulletin provisions would improve current agency practices.

The perchlorate example illustrates how federal agency risk assessments often give rise to disputes between and among federal agencies. As the agency charged with overseeing the management of the Executive Branch, OMB is uniquely suited to mediate such disputes, and so it is particularly appropriate for OMB to issue the proposed Bulletin. OMB also has historically provided agency guidance on risk analysis, in Democratic as well as Republican administrations.[5] While OMB has added internal staff capability in the areas of toxicology and epidemiology, OMB is also right to have consulted with the Office of Science and Technology Policy in developing the Bulletin and in proposing to continue that consultation in implementing it.

---

[3] 2003 Perchlorate State-of-theScience Symposium, Univ. of Nebraska Medical Center. http://www.perchloratesymposium.com/ and http://www.perchloratesymposium.com/modlist.asp. NRC, HEALTH IMPLICATIONS OF PERCHLORATE INGESTION (2005), *available at*: http://fermat.nap.edu/catalog/11202.html#toc. USEPA IRIS Perchlorate (ClO_4^-) and Perchlorate Salts: (last revised Feb. 18, 2005), *available at*: http://www.epa.gov/iris/subst/1007.htm.

[4] *Graham Cites Perchlorate Dispute To Defend Need For OMB Risk Guide*, RISK POLICY REPORT (May 30, 2006).

[5] *E.g*, OMB, Principles for Risk Analysis (Jan. 12, 1995).

B. The Bulletin Condenses a Quarter Century of Consensus Recommendations Regarding Risk Assessment Theory and Practice

The draft Bulletin can be best characterized as progressive "codification" of over 25 years of advice from august groups. Implementation of the Bulletin will help to capture the value of tremendous advances that have been made in risk assessment theory and practice since the publication of the NRC's "Red Book" in 1983.[6] As the Bulletin documents, since that date the National Academies, a Presidential/Congressional commission, EPA's Science Advisory Board and other prominent, authoritative bodies have fleshed out and advanced the state of the art in health and safety risk assessment. The Bulletin very helpfully crystallizes that "art" into a concise set of goals and expectations that federal agencies can consistently follow. Virtually all of its provisions reflect concepts that are well-established in deliberations and publications of the mainstream risk assessment community, a point emphasized at various places below.

C. OMB Has Ample Legal Authority to Issue the Bulletin

The Council believes that the provisions of the Bulletin are entirely within OMB's legal authority. Agency risk assessments made available to the public are "information" "disseminated" by those agencies, and hence fall within the scope of the IQA.[7] The provisions of the Bulletin are plainly intended to improve the quality of those assessments, principally by increasing their objectivity and utility, and so the Bulletin does further the purpose of the IQA and OMB's Guidelines under it. The Bulletin also constitutes OMB "guidelines . . . apply[ing] to Federal agency dissemination of public information" as authorized by the Paperwork Reduction Act.[8] ACC also agrees with the appropriateness of the Bulletin's citations to the Regulatory Right-to-Know Act[9] and Executive Order 12866.[10]

D. The Bulletin Does Not Impair Agencies' Ability to Implement Their Statutory Mandates

1. The Bulletin does not override any Congressional directions

Critics of the Bulletin have expressed concerns that the Bulletin "conflicts with statutory directives" and, by requiring agencies to consider factors such as countervailing risks, would prevent them from following the protective approach embodied in the nation's environmental and worker safety laws. These concerns are groundless. It is an

[6] NRC, RISK ASSESSMENT IN THE FEDERAL GOVERNMENT: MANAGING THE PROCESS (1983).
[7] 44 U.S.C. § 3516 note.
[8] *Id*. § 3504(d)(1).
[9] 31 U.S.C. § 1105 note.
[10] 58 Fed. Reg. 51735 (Oct. 4, 1993)

established principle of administrative law that OMB has no power to force agencies to violate the mandates those statutes create,[11] and the Bulletin does not suggest otherwise.

It is also true, however, that federal environmental and workplace safety laws almost uniformly address risk management – that is, they require or authorize various sorts of controls or limits at some given level of risk. For example, the Clean Air Act requires national primary ambient air quality standards for criteria pollutants to be set at a level that is "requisite to protect the public health" with an "adequate margin of safety."[12] The Bulletin properly states that it "does not encompass how federal agencies should manage . . . risk."

Very few federal statutes, however, address the prior question of how those risks are to be characterized or estimated, *i.e.*, how they are assessed. Indeed, the only statute of which ACC is aware that addresses risk assessment is the Safe Drinking Water Act – and that statute's requirements are actually incorporated into OMB's IQA Guidelines and the Bulletin, as discussed below in Part III.A. Consequently, the Bulletin addresses a topic to which Congress generally has not spoken – and on which constructive guidance would be very beneficial.[13]

Critics have also asserted that the Bulletin prohibits: (i) consideration of individual risks (allegedly, by requiring the reporting of population risks); and (ii) consideration of susceptible groups or individuals (allegedly, by requiring the reporting of central estimates and ranges of risk) – supposedly thus preventing agencies from following statutory directions to consider the risk to such entities. Neither assertion has merit. The Bulletin's requirements are deliberately flexible. Section IV.2.b states: "Each agency risk assessment shall . . . clearly summarize the scope of the assessment, including a description of . . . the affected entities (*population(s), subpopulation(s), individuals* . . .) that are the subject of the assessment."[14] Section IV.7.d adds: "For risk assessments that will be used for regulatory analysis, the risk assessment . . . shall include . . .estimates of population risk *when estimates of individual risk are developed*."[15] These passages make it abundantly clear that the Bulletin does not limit any agency's discretion in estimating individual risks or risks to highly-exposed or extra-susceptible groups or individuals. Rather, the Bulletin sensibly requires that risks be fully characterized in as transparent a manner as possible. Furthermore, Section V.3 will drive influential risk assessments to

[11] *See, e.g., EDF v. Thomas*, 627 F. Supp. 566 (D.D.C 1986). *See also* E.O. 12866, *supra* note 10, § 9. *See generally* American Bar Association Section of Administrative Law & Regulatory Practice, A GUIDE TO JUDICIAL AND POLITICAL REVIEW OF FEDERAL AGENCIES 227-228 (2005).
[12] 42 U.S.C. § 7409(b)(1).
[13] Admittedly, the way a risk is estimated could affect whether and how it is managed under a particular statute. However, federal environmental and workplace safety usually do not draw hard links between particular numerical expressions of risk (*e.g.*, 1 x 10^{-5} cancer risk) and particular management steps. Instead, they typically use vague verbal expression like "safe." Thus, regardless of how the Bulletin directs agencies' risk assessment efforts, those agencies generally retain the policy discretion to choose what estimate of risk equates to the words of a statute that trigger action.
[14] Bulletin at 24 (emphasis added).
[15] Bulletin at 25 (emphasis added).

present the full range of risks – from high end to central tendency to low end. This will provide a more comprehensive portrayal of full breadth and spectrum of the scientific data and therefore will be more informative to the risk manager.

2. The Bulletin will not ossify the regulatory process

Opponents of the Bulletin also contend that it imposes unreasonable additional burdens on agencies and will overload their capacities. To the contrary, the Council believes that:

The increased workload should not normally be overly burdensome. In fact, the Bulletin's requirements are intended to capitalize on the enormous scientific advances of the past three decades, and any increased workload would be a relatively small incremental increase over effort already expended. Because current practices are often times inadequate, the benefits of such an incremental increase in workload would likely be offset by time saved later in the process. Currently, many assessments require several, time consuming, iterations to achieve the required and necessary degree of comprehensiveness and objectivity. With implementation of the Bulletin, time spent by agencies on such iterations either would no longer be necessary, or at minimum would be reduced...

The Bulletin could actually conserve agency resources by promoting more reliable and defensible risk assessments. In comments before the National Research Council (May 22, 2006) and at a briefing for Congressional staff (May 24, 2006; sponsored by the American Chemical Society and Society for Risk Analysis), Dr. Graham asserted that delays in regulatory decision-making and waste of limited resources could be reduced or avoided if agencies acted on the Bulletin's requirements for promoting high scientific quality from the outset. Dr. Graham noted the requirement in the Bulletin that "assessments shall be a product of an iterative dialogue between the assessors and the agency decision makers."[16] He specifically pointed to the assessment that U.S. EPA conducted for human health risks from perchlorate exposure as an example of when and how more expeditious and effective regulation could have resulted from "getting it right the first time."

This perspective is derived directly from the NRC's recommendation in *Understanding Risk*, which focused on the need for deliberation between assessors and policy makers and the virtue of "[g]etting the science right."[17] *Understanding Risk* expressly dealt with the argument that these steps cost too much, but responded that "analyses . . . that do not . . . use reasonable assumptions . . . can result in huge expenses and long delays and jeopardize the quality of understanding and the acceptability of the final decisions.[18]

[16] Bulletin at 23.
[17] NRC, UNDERSTANDING RISK: INFORMING DECISIONS IN A DEMOCRATIC SOCIETY 3-7 (1996).
[18] *Id*. at 9-10.

Current agency risk assessment practice consumes significant resources, yet too often seems to follow the motto "Never enough time to do it right; always enough time to do it over." Risk assessments that are not scientifically or technically sound often have to be done over, as a result of either (i) compelling information and arguments supplied in the notice and comment process or (ii) judicial action based on such information and arguments. High-quality risk assessments are more likely to withstand both such challenges.

To the extent the Bulletin is more demanding than current practice, those demands are, in fact, justified. ACC believes that federal agency risk assessment practices, by and large, have not kept up with the state of the art in risk assessment as articulated in the Bulletin. In particular, they are failing to capture the enormous value of scientific progress (for example, in our understanding of mechanisms of toxicity). Given the need for risk-based agency decisions to be rational, fair and cost-effective, and the potentially enormous costs associated with such decisions, society can justifiably demand that agencies adopt the practices specified in the Bulletin.

II. Scope & Applicability Issues

A. The Bulletin Appropriately Applies to Documents that Are Components of Risk Assessments

Some speakers at the NRC conference argued that the Bulletin should only apply to complete risk assessments and not to dose-response or exposure assessments. ACC strongly disagrees. These sorts of component documents can be hugely influential. A prime example is EPA's IRIS assessments. Although IRIS assessments address only hazard identification and dose-response, and thus are not complete risk assessments, IRIS values (RfDs, RfCs, Cancer Potency Slope values) are widely used locally, at the state level, nationally and internationally as the toxicity characterization portion of site-, situation-, and media-specific risk assessments. IRIS values are routinely used in Superfund, air toxics and drinking water risk assessments. Another important example is the National Toxicology Program's technical reports on toxicity and Reports on Carcinogens, which, like IRIS values, are not complete risk assessments but comprise a critical component of a risk assessment.

It makes no sense to exclude such components of a risk assessment from compliance with the Bulletin. It would also be exceedingly inefficient to defer compliance with the goals and standards of the Bulletin until such time as these components are employed in a formal risk assessment. Instead, the Bulletin wisely makes clear that components of risk assessments, developed as stand alone evaluations with the intent of being widely used in site-, situation-, or media-specific risk assessments, must comply with the Bulletin. Obviously, they would only need to comply with the relevant provisions. Provisions

related to exposure assessment, for example, would not apply to a hazard characterization component.[19]

B. OMB Should Clarify that the Bulletin is Mandatory

The Council is encouraged by OMB's characterization of the draft Bulletin as "provid[ing] clear, *minimum* standards for the scientific quality of federal agency risk assessments."[20] The important substance of this characterization is repeated in the text of the draft Bulletin: "The purpose of this Bulletin is to enhance the technical quality and objectivity of risk assessments prepared by federal agencies by establishing *uniform, minimum* standards."[21] However, the very first words of the "applicability" section of the Bulletin are: "*To the extent appropriate*, all agency risk assessments…shall comply with the standards of this Bulletin."[22] Similar diction appears frequently throughout the document. The Bulletin's liberal use of non-directive language raises serious concerns that agencies will feel free to comply with it whenever, and to whatever extent, they choose. Such a result obviously would vitiate the Bulletin. In the comments that follow, the Council offers multiple recommendations for greater clarity and more directive tone.

C. OMB Should Clarify that the Bulletin Applies to Independently-Conducted Risk Assessments When Relied upon by Federal Agencies

As OMB has explained, the Bulletin is premised in part on OMB's legal authority under the IQA and "builds on" that act.[23] Under OMB's IQA Guidelines, information generated by persons other than the government is nonetheless subject to the IQA when the government disseminates it, whether by publishing it in some fashion or by relying on it in making a decision.[24] Accordingly, a risk assessment prepared by an entity other than the government should likewise be subject to the Bulletin whenever a federal agency disseminates it.

OMB should clarify this point and should effectively communicate it to providers of risk assessment services who are not otherwise covered by the Bulletin (e.g., universities,

[19] Provisions applicable to hazard identification and dose-response components, such as IRIS, would include, but not necessarily be limited to: problem formulation, completeness, effort, resources, peer review, public participation, standards related to informational needs and objectives, scope, objectivity, critical assumptions, executive summary, a portion of the standards related to risk characterization and regulatory analysis (specifically, those portions that call for a range of plausible estimates, including a central estimate), reproducibility, comparison to other results, comparison of numerical estimates, a portion of the standards for characterizing uncertainty (dose-response modeling and influence of assumptions), standards for characterizing results, etc.
[20] Press release, "OMB Requests Peer Review of Proposed Risk Assessment Bulletin" (Jan. 9, 2006) (emphasis added).
[21] Bulletin at 3 (emphasis added).
[22] Bulletin § II.1 (emphasis added).
[23] *Id*. at 3, 7.
[24] OMB, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*, 67 Fed. Reg. 8452, 8454 (Feb. 22, 2002).

contractors, consultants, etc.). Doing so would alert those providers that if federal agencies rely on their independently-developed work product, it will be subject to the criteria in the Bulletin (as well as the Peer Review Bulletin and the other standards of the IQA).

D. The Bulletin Should Apply to Permitting and Licensing Decisions

The Council believes that OMB should eliminate the exclusion for adjudication and permitting (Section II.2.b), at least in cases where an agency is both developing a risk assessment and making a decision based upon it (as is usually the case). There is already substantial precedent for risk assessments that are done for "adjudicative" processes (e.g., a pesticide registration under FIFRA) to establish science policies and practices that are then exported to other applications and situations.[25] As presently drafted, the exclusion for such risk assessments would substantially undercut the Bulletin's goal of establishing uniform standards for federal risk assessments. The exclusion also provides fodder for critics of the Bulletin, who contend that the exclusion allows some of the most important risk assessments to avoid the Bulletin's standards.[26]

The Council understands that the exclusion is a manifestation of OMB's general practice of addressing agency decisions of general, rather than particular, applicability.[27] However, risk assessments in many agency permitting and licensing contexts can have precedential effect outside of those contexts, and thus effectively are of general applicability.[28]

We also recognize that a similar exclusion is contained in the Peer Review Bulletin. However, the considerations possibly supporting the exclusion in that situation do not apply here. While a permit applicant may be able to arrange for a peer review of an agency assessment, permit applicants certainly cannot assure the initial quality of that assessment. The Bulletin is needed to ensure that agencies conduct high quality assessments in such cases. Elimination of the exclusion will hold agencies to the

---

[25] For example, in 1988, EPA's pesticide program office reviewed the existing science on thyroid follicular cell carcinogenesis and developed a science policy position covering the evaluation of chemicals that have induced thyroid tumors in experimental animals, which concluded that human risks should be assessed using nonlinear models. This approach was then approved by the Agency's FIFRA Scientific Advisory Panel, and has subsequently been applied to such agents across EPA program offices. http://www.epa.gov/pesticides/trac/science/cumulative_guidance.pdf

[26] Statement of Rena Steinzor, University of Maryland, before Congressional Research Service Conference on Science in Rulemaking, American University, Washington, DC (May 9, 2006), *available at* www.american.edu/rulemaking/ppt/0510rulemaking.doc (*see* transcript at page 13).

[27] Statement of Donald Arbuckle, Acting OIRA Administrator, before Congressional Research Service Conference on Science in Rulemaking, American University, Washington, DC (May 9, 2006), *available at* www.american.edu/rulemaking/ppt/0510rulemaking.doc (*see* transcript at page 18); *cf.* E.O. 12866, § 3(d), (e).

[28] For example, in a Superfund or RCRA action, an Agency assessment using exposure parameters and toxicity criteria to set soil cleanup levels may set a precedent for risk assessment methods or environmental levels at non RCRA/Superfund sites

Bulletin's standards, thereby protecting the interests of pesticide registrants, Resource Conservation and Recovery Act (RCRA) permit applicants, Superfund potentially responsible parties, etc. Furthermore, elimination of the exclusion is consistent with the stated goals of the Bulletin, i.e., assuring that the best available science is applied to guide regulatory decision-making.

ACC therefore urges OMB to eliminate the exclusion in cases where the agency is both the (i) submitter of the risk assessment and (ii) the decision-maker. This approach, proposed by the American Bar Association,[29] would address the cases of concern to ACC, and would also retain the existing exclusion for other situations that are more quintessentially adjudicatory. ACC also believes OMB should shift the burden of proof, as it were, so that the Bulletin applies *unless* an agency makes certain findings. The modified text of Section II.2.b should read:

> b. individual agency adjudications or permit proceedings <u>where the agency is not both the submitter of the risk assessment and the decision-maker, when</u> the agency determines that:
>
> i. compliance with the Bulletin is <u>im</u>practical and <u>in</u>appropriate and
>
> ii. the risk assessment is <u>neither</u> scientifically <u>nor</u> technically novel, <u>nor</u> likely to have precedent-setting influence on future <u>risk assessments or important public policies or private sector decisions</u>.

E. The Bulletin Should Not Allow for Waivers, Only Deferrals

The draft Bulletin allows for both waivers and deferrals. The Council agrees that agency heads should have the option to defer compliance with the provisions of the Bulletin in rare cases where exigent circumstances truly demand expedited action. However, the Council also believes that agencies should be required to complete a quality assessment, as otherwise required by the Bulletin, when those circumstances have passed. Simply put, there is no basis for perpetually exempting a decision simply because it was made during an emergency. Indeed, such decisions are the ones most likely to be flawed and most likely to benefit from application of the Bulletin.

Allowing only deferrals would provide two benefits. First, an agency may find that the expedited measures (taken under hurried circumstances) are either insufficient or excessive. With better risk information in hand, the agency can correct the situation by revisiting the measures taken. Second, applying the Bulletin in due course would ensure that decisions made under hurried circumstances do not establish precedents that could be misapplied in other, less demanding circumstances.

[29] American Bar Association, Comments on OMB's Proposed Risk Assessment Bulletin 8-9 (May 22, 2006).

F. The Bulletin should clarify the meaning of "regulatory analysis"

Section IV.7 introduces a new concept, "regulatory analysis." In her remarks before the National Academy of Sciences meeting on May 22 and the Society for Risk Analysis symposium on May 23, Dr. Nancy Beck (OIRA,OMB) stated that the term includes any analysis conducted in support of a rulemaking, and is not limited to regulatory impact analyses under E.O. 12866, regulatory flexibility analyses under the Regulatory Flexibility Act, and the like. The Council supports this broad definition and recommends that the Bulletin specifically clarify this point.

G. A More Comprehensive Inventory of "Influential" Risk Assessments Is Needed.

The Council supports the Bulletin's use of the same definition of "influential" as is used in OMB's IQA Guidelines and the Peer Review Bulletin. The Council also supports the principle behind OMB's listing of "examples of influential risk assessments" in the preamble of the Bulletin. And we support OMB's explicit statement that the list is non-exclusive. However, numerous other assessments are often cited, and can have significant influence, beyond the domain of their original development. Among those are:

- various health or environmental criteria,
- health (effects) advisories,
- provisional assessments (as prepared for or by the Superfund office of EPA), and
- special assessments by other offices or agencies that are known by various names (e.g., "health reference levels," as referred to on the Web site of EPA's Office of Air Quality Planning and Standards).[30]

The Council is especially concerned about the quality of HEAST (Health Effects Assessment Summary Tables). HEAST values, including tabular information on such key risk-related statistics as "carcinogenicity slope factors," are frequently cited as authoritative health risk information.[31] HEAST is a secondary (or sometimes tertiary) source and therefore of questionable integrity. It also does not appear to have been peer-reviewed or otherwise quality-assured in any rigorous way, and the HEAST values can be considerably out of date.

The Bulletin acknowledges that "a risk assessment prepared by one federal agency may inform the policy decisions of another federal agency."[32] It also recognizes that federal agency assessments "can directly or indirectly influence the regulatory actions of state

[30] See http://www.epa.gov/oar/oaqps/air_risc/3_90_024.html.
[31] See http://www.epa.gov/radiation/heast/. HEAST is also cited in the EPA's Soil Screening Guidance (SSG) – the SSG "presents a framework for developing risk-based, soil screening levels (SSLs) for protection of human health."http://www.epa.gov/superfund/resources/soil/
[32] Bulletin at 3.

and local agencies or international bodies."[33] Indeed, due to the limited availability of resources, a multitude of regulatory authorities outside the United States rely very heavily on U.S. agency-generated assessments as reliable sources of information on risk and safety.

OMB is also clearly aware of this potential for the application of risk assessments well beyond the context in which those assessments were developed -- and well beyond the legitimate application for which the assessments were intended.

Consequently, the Council recommends that OMB conduct an inventory of federal agency risk assessments that should be regarded as influential. That inventory should then be employed to expand the list of examples to more completely enumerate those assessments that are subject to Section V of the Bulletin.

H. The Bulletin Should Reinforce the Applicability of "Minimum Standards" to Ecological Risk Assessment

The draft Bulletin provides only limited specific guidance to assure the quality of the assessments for threats to non-human targets or assets (i.e., ecological risk assessments). While the provisions of the Endangered Species Act, the Marine Mammal Protection Act, Section 404 of the Clean Water Act and similar statutes present unique challenges that the Bulletin does not directly address, OMB should reiterate that the "minimum standards" for quality, utility and integrity apply to <u>all</u> assessments, including those for ecological risks.

Given the urgency for issuance, the Council recommends that OMB move ahead in issuing the Bulletin as it applies to human health risk/safety assessments. OMB should encourage efforts among affected agencies to advance appropriate guidance to deal with the special challenges of ecological assessments. In the meantime, the Council is eager to partner with OMB, other stakeholders, and the relevant professional communities in a consensus-building exercise to fill critical gaps.

III. The Bulletin Should Have Clearer and Stronger Linkage to the Information Quality Act and Associated OMB Guidelines

The Council has long argued that the Information Quality Act (IQA) is a vital supplement to the Administrative Procedure Act because it assures that federal agencies are accountable for the power they can exercise through the dissemination of information. ACC has been an unparalleled proponent of OMB's efforts to implement the IQA, including OMB's IQA Guidelines, Peer Review Bulletin, and memoranda to agencies. The Risk Assessment Bulletin "builds on" IQA guidelines constructively. However, the Council also believes that the linkage between the Bulletin and the IQA could be more clearly and strongly established, as discussed below.

[33] Bulletin at 9.

A. Safe Drinking Water Act Provisions

In the Council's opinion, the single most important provision of the Bulletin is Section IV.4, requiring that all assessments be "scientifically objective." That provision should have the effect of abolishing certain long-established, policy-biased practices that lead to overestimation of risks by orders of magnitude and resultant misallocation of resources.

- The requirement of Section IV.4.a that risk assessments "neither minimiz[e] nor exaggerat[e] the nature and magnitude of the risk" should have the effect of abolishing the practice of selecting data sets (e.g., animal test results) and risk estimation models with the specific purpose of maximizing the estimated risk:
    - In the case of animal test results, common agency practice is to assume humans are as sensitive as the most sensitive animal species/strain/response and therefore agencies select the data set that leads to the highest estimate of risk.
    - In the case of exposure assessment, common agency practice is to rely on "high end estimates."
    - In the case of selecting "points of departure" (PODs) and risk extrapolation models, common agency practice is to select those PODs or models that systematically yield the highest risk values (or lowest "safe dose" values).

These and some other practices can be inconsistent with Section IV.4.a, and ACC urges OMB to retain that provision unchanged. From the perspective of scientific objectivity, the Bulletin would lead to a fuller portrayal of the scientific foundations of a risk assessment. For example, the Bulletin will improve risk assessment by assuring that the relevance of animal toxicity results to humans be included in the deliberative process and that the impact of alternative, biologically plausible assumptions be examined quantitatively. Presentation of the distribution of exposures and risks, including central tendency and high-end, based on a comprehensive and objective evaluation of the most relevant data will provide a more complete and scientifically accurate representation of potential risks. In this regard, it is important to note, that the Bulletin does not impose any risk management criteria upon an agency. The improvements in risk assessment that will be brought about by the Bulletin will not preclude agency consideration of risk management decisions at some prescribed level, be that high-end, 90% or 95%, the most sensitive or susceptible or the at a more central metric. The Bulletin assures that the risk manager and the public are more fully informed about what is known scientifically, what is assumed, and how potential risks are distributed across the population. The Bulletin does not dictate that risk managers focus on the median, or the central tendency.

- The requirement of Section IV.4.b that risk assessments "giv[e] weight to both positive and negative studies" in light of each study's technical quality should have the effect of requiring that all evidence to be accounted for, leading to an assessment that is both comprehensive and objective. Agencies often follow a

practice of basing assessments on a single positive study, no matter the comparative quality or number of negative studies.

- ACC also strongly supports Section IV.3, which requires risk characterizations to be provided quantitatively whenever possible, and for such quantitative expressions to include a range of plausible risk estimates. Agency risk assessments too often have avoided quantifying risks, or have presented only a single point estimate.

The preamble to the Bulletin explains that, beyond complying with the literal words of Sections IV.3 and IV.4, risk assessments must also comply with the language from the Safe Drinking Water Act (SDWA) regarding, respectively, presentational and substantive objectivity. The Council strongly supports this mandate. As noted in Part I.D.1 above, the SDWA language is one of the few instances in which Congress has actually directed the way in which federal agencies conduct and present risk assessments. These provisions are a core element of OMB's IQA Guidelines. The Council makes two additional recommendations in this connection:

First, in order to more firmly establish this important point, the Council recommends that the relevant language from the SDWA (42 U.S.C. § 300g-1(b)(3)(A) & (B)) be explicitly included in the actual text of Sections IV.3 and IV.4 of the Bulletin.

Second, OMB's IQA Guidelines currently apply the SDWA requirements only to *influential* risk assessments.[34] The Risk Assessment Bulletin clearly expands that requirement to "*all* risk assessments subject to this Bulletin."[35] ACC strongly supports this expansion.

Additional technical comments on Sections IV.3 and IV.4 are contained in Part VII.B below.

B. The Peer Review Bulletin Is Still Mandatory

ACC has actively supported OMB's Peer Review Bulletin, which applies to agencies "[t]o the extent permitted by law."[36] Consistently, the Risk Assessment Bulletin's Goals section (Section III.5) states that each "agency shall follow appropriate procedures for peer review." ACC is concerned, however, by the preambular discussion of "peer review and public participation," which says that agencies "should consider appropriate procedures" for peer review.[37] OMB should revise the preamble to make it consistent with the Bulletin itself, and should clarify that the provisions of the Peer Review Bulletin are binding according to its terms.

[34] See Guidelines Section V.3.b.ii.C, 67 Fed. Reg. 8460.
[35] Bulletin at 14 (emphasis added).
[36] OMB, *Final Information Quality Bulletin for Peer Review*, § II.1, 70 Fed. Reg. 2664, 2675 (Jan. 14, 2005).
[37] Bulletin at 9.

C. ALL risk assessments must be "capable of being substantially reproduced."

The Bulletin states that all "[i]nfluential risk assessments should be capable of being substantially reproduced."[38] While we recognize that this position is consistent with OMB's IQA Guidelines regarding the reproducibility of scientific, financial and technical information generally, the Council believes that limiting the reproducibility requirement to influential risk assessments is unwise. *All* scientific assessments used in decision making, regardless of scope, scale, depth, detail, or purpose, should be reproducible.

The Council's position is predicated on the common-sense reality that assessments that cannot be reproduced are valueless. Indeed, irreproducible assessments actually have negative value -- they convey information of such low quality and untrustworthy character that they can easily contribute to ill-advised decisions. The common-sense point may be best captured in the rhetorical question, "what value has an assessment if 'reanalysis of the original or supporting data, using the same methods' leads to inexplicably different results?" An irreproducible assessment has no legitimacy in science and should have no place in policy-making.

Therefore, the Council strongly urges OMB to revise the Bulletin to provide that all scientific assessments must be "substantially reproducible."

D. The Bulletin Needs to Emphasize the Critical Nature of Improved Risk Assessments to Inform Benefit Cost Analyses

Risk assessments are an integral part of benefit-cost analysis (BCA), as they are necessary to evaluate the benefits of various courses of action (i.e., what risks will be reduced and by how much?). BCA, in turn, is crucial to full and effective implementation of President Clinton's Executive Order 12866 and Circular A-4[39]. OMB should more clearly and compellingly link the need for advancements in risk assessment with the need to improve benefit-cost analyses, regulatory impact analyses, and final rulemaking.

In particular, the Council urges OMB to cite the relevant passages from the Executive Order to assure that agencies fully appreciate their obligation to comprehensively assess risks with respect to regulatory objectives and regulatory alternatives. The Council recommends that the Bulletin be revised to cite the following important provisions of E.O. 12866:

> Section 1.a. In deciding whether and how to regulate, agencies should assess all costs and benefits of available regulatory alternatives, including the alternative of not regulating."

[38] Bulletin at 16.
[39] OMB, *Circular A-4* (Sept. 2003).

> Section 1.b.5. When an agency determines that a regulation is the best available method of achieving the regulatory objective, it shall design its regulations in the most cost-effective manner to achieve the regulatory objective.

(Emphasis added.)

IV. Public Participation Should Be Required for all Risk Assessments

The Bulletin should establish a general expectation that agencies will seek public participation in risk assessment. The Council believes that public participation – early, meaningful and at appropriate stages in the process – is beneficial, and that special efforts should be expended to secure the benefit. The National Research Council, the American Bar Association and other authoritative bodies have likewise highlighted the importance of involving interested parties early and meaningfully in scoping and conducting risk assessments, both to enhance their credibility to those parties and to ensure that they are objective and based on the best available information.[40]

As in the case of peer review, the Bulletin and the preamble are inconsistent on this point. The Bulletin itself directs that agencies "shall follow appropriate procedures for…public participation."[41] However, the preambular discussion of "peer review and public participation" states only that agencies "should consider appropriate procedures for…public participation."[42]

ACC respectfully submits that neither provision is quite right. ACC recommends instead that OMB revise both the Bulletin and its preamble to say simply that agencies shall enable public participation in all risk assessments. At a minimum, the Bulletin should require public notice of influential risk assessments, accompanied by a solicitation of comments. The notice should be published early enough, and allow time for thoughtful input at the scoping/formulation stage and throughout the assessment process.

ACC also notes the inconsistency between the less-than-mandatory language regarding public participation and the mandatory language regarding consideration and response to comments. As to the latter, the Bulletin states that agencies "shall consider all significant comments received on a draft influential risk assessment"[43] and "shall issue a 'response-to-comment' document."[44] Such mandatory language may lead agencies to look for ways not to seek public comments in the first place. The solution is to make the initial solicitation mandatory as well.

---

[40] *See, e.g.*, UNDERSTANDING RISK, *supra* note 17, at 23-24, 73-96; American Bar Association, Resolution on Risk Assessment (Aug. 1999); *available at*: http://www.abanet.org/adminlaw/risk02.pdf.
[41] Section III.5.
[42] Bulletin at 11.
[43] Bulletin at 25.
[44] Bulletin at 25.

V. OMB Should Prescribe Implementation Procedures for the Bulletin

The Council supports the spirit of Section IX, OIRA and OSTP Responsibilities. However, the Council believes that in order to achieve the Bulletin's ambitious goals, OMB (in cooperation with OSTP) must be clearer on exactly how it will monitor and oversee agency implementation of the Bulletin. The IQA Guidelines and the Peer Review Bulletin both established clear procedures for agencies to follow and explained how OMB would oversee them. The Risk Assessment Bulletin is silent on this topic, however.

The comments of the Coalition for Effective Environmental Information, to which ACC belongs, focus almost exclusively on this issue. In particular, they propose that OMB require agencies to publish implementation plans that explain how they will revise their existing risk assessment policies in light of the Bulletin's requirements. ACC supports those comments and urges OMB to follow their recommendations.

VI. Additional Administrative Guidance is Needed

A. Maximum "Return on Investment" in Research and Testing Requires that Influential Risk Assessments Be Regularly Updated, and that Updated Risk Assessments Be Disseminated in Lieu of Outdated Ones.

Both private- and public-sector parties invest substantial resources in research and testing to support risk assessment. Sponsors and taxpayers alike should surely expect a "fair return" on their investments, i.e., that the outputs from their research and testing do, in fact, advance risk assessment and regulatory decision-making. The Council believes that the substantial nature of investments in research and testing demand that influential assessments be regularly updated.

Authoritative bodies have recognized the need for ongoing updates. Both the NRC and the Presidential/Congressional Commission on Risk Assessment and Risk Management (Risk Commission) have pointed out that older assessments are likely to have been conducted using "science policy defaults" (a.k.a. default inference options), and that, as the fund of scientific knowledge increases, the inclination to rely on defaults should diminish.[45]

[45] National Research Council, *Science and Judgment in Risk Assessment* 90 (1994) ("Over time, the choice of defaults should have decreasing impact on regulatory decision- making. As scientific knowledge increases, uncertainty diminishes. Better data and increased understanding of biological mechanisms should enable risk assessments that are less dependent on default assumptions and more accurate as predictions of human risk."); 2 Risk Commission, RISK ASSESSMENT AND RISK MANAGEMENT IN REGULATORY DECISION-MAKING iv (1997) ("Agencies should continue to move away from the hypothetical . . . toward more realistic assumptions based on available scientific data.").

Also, there is no point in updating risk assessments if agencies continue to disseminate the outdated assessments they replace. For example, a critical Web site from EPA's Office of Solid Waste and Emergency Response (OSWER) does not include an internet link to the most current – and presently applicable – Agency-wide guidelines for assessing cancer risks. In April 2005, EPA issued its final cancer risk assessment guidelines.[46] Despite the availability of those guidelines for more than a year, no reference to them appears on the OSWER Web site.[47] The final Bulletin should mandate that agencies promptly disseminate updated risk assessments and clarify that any older documents that it disseminates are provided for historical purposes only.

In the interest of capturing the value of society's investment in risk assessment, as well as in pursuing the declared goals of the Bulletin (i.e., improving the technical quality and objectivity of federal risk assessments), the Council recommends revising Section VI of the Bulletin to read as follows:

> As relevant and scientifically plausible information becomes available, each agency shall update and replace its assumptions and assessments to reflect new data or scientific understandings."[48] [emphasis added to highlight changes]

OMB should specifically state that, as a result of being updated, existing risk estimates may either remain unchanged or be revised downward or upward based on new scientific information.

B. Comparison with Other or Previously Conducted Assessments on the Same Topic

It is reasonable to expect that implementation of the Bulletin will lead to improved risk assessments outside the federal government as well as within it. In particular, the Council suggests that Section V.2 may foster a scientific environment in which interested parties are encouraged to compete to develop high quality assessments (i.e., assessments that meet or exceed the requirements of the Information Quality Act, OMB Information Quality Guidelines, the OMB Peer Review Bulletin, and this draft Bulletin).

The Council believes that the Bulletin will foster scientific excellence at all agencies and provide assurance across agencies that the quality and scientific standards have been met. ACC has two suggestions to improve Section V.2.

First, the Bulletin should clarify that when agencies compare the results of their influential risk assessments to prior published risk assessments, the standards described in the Bulletin are paramount. As the Bulletin drives federal agency risk assessments to a new level of sophistication and scientific quality, it is important that agencies not be able

[46] Available at: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=116283.
[47] See http://www.epa.gov/oswer/riskassessment/human_health_toxicity.htm.
[48] Bulletin at 25, with recommended changes noted

to disregard their own findings in deference to previously conducted assessments that do not meet similar standards. Thus, when comparing to previously conducted risk assessments, agencies should include a discussion section describing how the previous assessment meets, or falls short of meeting, the Bulletin's standards.

Second, it is important that agencies not be able to undermine the goals of the Bulletin by delegating activities to parties that are not subject to the Bulletin, i.e., through risk assessments that they conduct by funding second parties or otherwise coordinating with ostensibly independent third parties. Allowing such comparisons could create an incentive for agencies to generate inconsistency.

C. Screening-Level Risk Assessments Can Be Used Only to Support Decisions to Determine Whether to Regulate

The preamble to the Bulletin states that, "if an agency is only interested in a screening-level assessment, then an assessment which explores alternative dose-response models may not be warranted."[49] As a reasoned principle of resource conservation, this statement is appropriate. But clearer guidance is required to assure that agencies use screening-level assessments only to identify whether circumstances warrant further attention. Agencies should never proceed to regulation without completing the scope and scale of assessment that is specified in the draft Bulletin (i.e., satisfying the "minimum standards for the scientific quality of federal agency risk assessments").

The Council agrees that unnecessary impediments to timely rulemaking are to be avoided. However, the Council also believes that the unqualified exemption that is extended to "screening assessments" (see page 9, *Section II: Applicability*) is ill-advised.

> "Section II states that, *to the extent appropriate*, all publicly available agency risk assessments shall comply with the standards of this Bulletin. This statement recognizes that there may be situations in which it is not appropriate for a particular risk assessment to comport with one or more specific standards contained in this Bulletin, including the general standards in Section IV, which apply to both influential and non-influential risk assessments."[50] [emphasis in the original]

The Council recommends that OMB require more complete characterization of risks in all circumstances, even those for which the only intended purpose is screening, such as in "priority setting." Even simplified assessments must include estimates of central tendency and discussion of uncertainties.

As previously noted, the potential for mis-applications (whether intentional or not) of screening-level assessments is very real. Despite all disclaimers, screening level toxicity

[49] Bulletin at 11.
[50] Bulletin at 9.

values can be mis-applied by federal, state, regional or even international offices or agencies. Given such ample opportunities for mis-application, the Council recommends that OMB require inclusion of central tendency estimates of risk (in addition to estimates that are based on "conservative (or worst-case) assumptions").[51]

There are at least two compelling reasons for requiring that "screening assessments" be based on a more complete assessment of available information. The first is that policy-neutrality demands it. That principle must apply to all assessments…and most certainly to include those that are applied to exclude risks from further consideration. A screening-level risk assessment which indicates low risk and thereby leads to a judgment that no (further) action is necessary is, in fact, a quasi-regulatory decision.. All assessments (including "screening assessments") on which any policy decision is predicated must be based on the most complete and best available evidence.

The second reason is related more to pragmatic matters of applying risk information in sensible ways. The upshot of systematically overestimating risks (for any purpose, however limited or extensive) is to destroy much of the utility of the assessment data. For this particular example (i.e., "screening risks"), consider the situation in which a set of risks are being compared for policy priority setting. If all the candidate risks are systematically over-estimated (by some variable, but uncertain amount), the value of the priority setting exercise is compromised. The utility of the relative "ranking" of risks is diminished by systematic and deliberate over-estimation.

D. Additional Guidance is Needed on How to Make Risk Assessments "Commensurate" with Risks

The Council agrees with OMB's position that investments in risk assessments (e.g., scope, scale, depth and detail) must be commensurate with the expected magnitude or severity of the risk, the available data and the decision needs. Here, as elsewhere, OMB is helpfully codifying a key recommendation of the Risk Commission.[52] The Council notes that while the Bulletin itself uses the word "shall" in the relevant passage, the preamble inconsistently uses the word "should."[53] OMB should conform the preamble to the Bulletin.

---

[51] The benefits from this simple requirement are substantial. Requiring more complete, richer risk characterizations for screening level assessments increases the likelihood that readers, including risk managers, will have a better understanding of the range of plausible risks and information on the level of certainty/uncertainty in the values (albeit more limited than that required for a comprehensive risk assessment).

[52] *See* RISK COMMISSION REPORT, *supra* note 45, at 63 ("Deciding to go forward with a risk assessment is a risk management decision, and scaling the effort to the importance of the problem, with respect to scientific issues and regulatory impact, is crucial.") and 21 ("The level of detail considered in a risk assessment and included in the risk characterization should be commensurate with the problem's importance, expected health or environmental impact, expected economic or social impact, urgency, and level of controversy, as well as with the expected impact and cost of protective measures.").

[53] *Compare* the Bulletin, at page 23 ("The level of effort put into the risk assessment shall be commensurate with the importance of the risk assessment.") *with* the preamble, at page 11 ("The level of effort should be

Moreover, there does not appear to be any broad consensus on how to scale risk assessments "proportionally." Consequently, the Council is concerned that the laudable goal of proportional investment may go unrealized. In the interests of advancing that goal, the Council is committed to working with OMB and other interested parties to discuss what guidance might be developed.

VII. Additional Technical Guidance Is Needed

As general matters, the Council recommends that OMB revise the Bulletin to further strengthen its directives for technical excellence. The Bulletin frequently includes the phrase "…wherever (or whenever) possible (or feasible)…" A few key examples are offered in the comments that follow The Council recommends that throughout the Bulletin OMB substitute the phrase "…to the greatest extent feasible…" which is language taken directly from the National Research Council's report *Science and Judgment in Risk Assessment*.[54] The Council supports OMB's generous use of the directive "shall" in the Bulletin (proper), but recommends that the preamble be made more consistent with the compulsory "shall" by routinely substituting "must" for the frequently used "should."

A. The Bulletin Should Clarify the Meanings and Applications of "Adverse" and "Non-adverse"

The Council strongly supports the clearly stated requirement that judgments regarding whether a specific effect is adverse (or not) "shall be specifically identified and justified."[55] The Council notes that assessments of the potential for non-adverse effects (e.g., adaptive changes) have not been the basis for rulemaking in the past, and the Council recommends that non-adverse effects not become the basis for regulatory decision-making except and until their utility and reliability have been demonstrated.

By the same token, ACC notes that exactly how to determine what constitutes an "adverse" effect is an unresolved question. OMB should partner with experts drawn from government agencies, the private and non-governmental sectors and academic/research institutions to develop a science-based consensus on this important issue. At the same time, the Council finds the reference to "…relevant clinical and toxicological communities" to be undesirably vague.

Given the challenges for developing specific guidance on this matter that is widely applicable across the spectrum of biological, chemical and physical mechanisms of action, the Council recommends that the passage in the Bulletin be revised to read:

---

commensurate with the importance of the risk assessment, taking into consideration the nature of the potential hazard, the available data, and the decision needs.") (emphasis added).

[54] Bulletin at 13 note 26.

[55] Bulletin at 20

> "Where human health effects are a concern, determination of which effects are adverse shall be specifically identified and justified based on the best available scientific information."[56]

B. Stronger Provisions Are Needed to Assure Reliance on the Best-Available Information, Absence of Bias, and Full Transparency.

The Bulletin must direct compulsory application of "best available science" and "weight of evidence.[57]" When disagreements arise among conscientious experts as to what constitutes the "best available science" and how to "weigh the evidence," it is imperative that fully transparent and scrupulous documentation of the bases for the controversy be required. Notwithstanding controversy in specific cases, OMB must leave no doubt that the best available information and most contemporary methods must be applied. Support for the Council's position comes from the report of the Risk Assessment and Management Commission. [58] The Council recommends revising the draft Bulletin as follows:

> "Beyond the basic objectivity standards, risk assessments subject to this Bulletin must use the best available data and must be based on the weight of the available scientific evidence."[59] [emphasis added to highlight changes]

As previously stated, the Council supports OMB's application of the SDWA provisions and takes special note that the draft Bulletin stipulates applicability to "…all risk assessments which address adverse health effects." [emphasis added] However, we recommend that the ensuing statement of guidance (i.e., the sentence that follows immediately from the previous quotation) be revised in order to strengthen agencies' commitments to the SDWA model.

---

[56] Bulletin at 20, with recommended changes noted

[57] The Council draws attention to the fact that a multi-stakeholder, collaborative project is currently underway (with participation by universities, U.S. and Canadian agencies, and private sector sponsors; coordinated by the International Life Sciences Institute) to develop a science-based consensus on the best practices for "weight of evidence" determinations. *See also* D.L. Weed, *Weight of Evidence: A Review of Concept and Methods*, 25 RISK ANALYSIS 1545 (2005).

[58] 1 RISK COMMISSION REPORT at 38 ("Because so many judgments must be based on limited information, it is critical that all reliable information be considered. Risk assessors and economists are responsible for providing decision-makers with the best technical information available or reasonably attainable, including evaluations of the weight of the evidence that supports different assumptions and conclusions."). The Risk Commission Report provides examples of the kinds of considerations entailed in making judgments on the basis of the weight of the scientific evidence in a toxicity study: quality of the toxicity study; appropriateness of the toxicity study methods; consistency of results across studies; biological plausibility of statistical associations; and similarity of results to responses and effects in humans. 2 RISK COMMISSION REPORT at 20.

[59] Bulletin at 14, with recommended changes noted.

> "These SDWA quality standards must be met, to the greatest extent feasible, in all risk assessments which address adverse health effects."[60] [emphasis added to highlight changes]

The Council supports the explicit requirement that risk/safety assessments be "unbiased" (throughout most of these comments, the Council has used the term "policy neutral," which we regard as synonymous).[61]

It is essential to assuring transparency that agencies are compelled to present their assessments in detail…and to justify fully their choices and judgments. To that end, the Council recommends that the Bulletin be revised to read:

> "Results based on different effects observed and/or different studies must be presented…When relying on data from one study over others, the agency must discuss fully the scientific justification for its choice."[62] [emphasis added to highlight changes]

The Council supports the principles from *Section IV*, subsection *5. Standards Related to Critical Assumptions*, but recommends revision as follows:

> "Risk assessments must explain the basis of each critical assumption and those assumptions which affect the key findings of the risk assessment. If the assumption is supported by, or conflicts with, empirical data, that information must be discussed. This must include discussion of the range of scientific opinions regarding the likelihood of plausible alternate assumptions and the direction and magnitude of any resulting changes that might arise in the assessment due to changes in key assumptions. To the greatest extent feasible, a quantitative evaluation of reasonable alternative assumptions must be provided. If an assessment combines multiple assumptions, the basis and rationale for combining the assumptions must be clearly explained."[63] [emphasis added to highlight changes]

The following passage is taken from *Section V*, subsection *4. Standard for Characterizing Uncertainty* (and corresponds to the passage above, except that *Section V* applies to the more stringent case of "influential risk assessments"). The Council submits that revision is an even more urgent matter in this context. The Council recommends:

---

[60] Bulletin at 14, with recommended changes noted.
[61] "In addition to meeting substantive objectivity standards, risk assessments must be accurate, clear, complete and unbiased in the presentation of information about risk. The information must be presented in proper context. The agency also must identify the sources of the underlying information (consistent with confidentiality protections) and the supporting data and models, so that the public can judge for itself whether there may be some reason to question objectivity. Data should be accurately documented, and error sources affecting data quality should be identified and disclosed to users." *See* Bulletin at 14.
[62] Bulletin at 19, with recommended changes noted.
[63] Bulletin at 15, with recommended changes noted.

> "When one or more assumptions are used in a risk assessment, the assessor must evaluate how plausible changes in the assumptions influence the results of the assessment."[64] [emphasis added to highlight changes]

C. General Requirements for Full Risk Characterization Should Be Strengthened.

The Council strongly endorses OMB's positions with respect to full characterization of risk, but recommends further strengthening the provisions of the draft Bulletin with respect to the presentation of central estimates of risk, specifically by clarifying that such a presentation is a requirement. The Council notes the authoritative context for the requirement that risks be fully characterized, as specified in the Safe Drinking Water Act.

> The agency is…directed "in a document made available to the public in support of a regulation [to] specify, to the greatest extent feasible — …(ii) the expected risk or central estimate of risk for the specific populations [affected]; (iii) each appropriate upper-bound or lower-bound estimate of risk;"[65] [emphasis added]

The Council notes OMB's observation that respected authoritative bodies have consistently recommended such reforms.[66,67]

OMB clearly recognizes the value-added from requiring that both central and limit estimates of risk be reported. To be consistent with the statutory requirement from the Safe Drinking Water Act, as well as consistent with the wise recommendations from authoritative bodies, the Council believes that the last sentence in this subsection should read:

> "The practice of highlighting only high-end or only low-end estimates is not acceptable."[68] [emphasis added to highlight the change]

The Council believes that the interests of "…quality, objectivity, utility and integrity…" are best served if the draft Bulletin is revised to read:

> "Influential risk assessments should characterize uncertainty with a sensitivity analysis and, to the greatest extent feasible, through use of a numeric distribution (e.g., likelihood distribution of risk for a given individual, exposure/event scenario, population, or subpopulation)."[69] [emphasis added to highlight changes]

---

[64] Bulletin at 18, with recommended changes noted.
[65] Bulletin at 13.
[66] Bulletin at 13 note 26.
[67] Bulletin at 13 note 27.
[68] Bulletin at 17, with recommended changes noted.
[69] Bulletin at 17, with recommended changes noted.

In the draft Risk Assessment Bulletin ("proper"), *Section* IV, subsections *3.*, *5.*, and *6.* (see page 24), there are 3 insertions of the modifier "whenever possible." For brevity, only the first (of three) occurrences is presented here.

> "IV. General Risk Assessment and Reporting Standards. Each agency risk assessment shall:
> …
> 3. Provide a characterization of risk, qualitatively and, whenever possible, quantitatively. [emphasis added]

In this and the two ensuing passages, the Council recommends substituting the phrase, "to the greatest extent feasible."

The Council strongly agrees with OMB that characterization of uncertainty must be mandatory for all influential risk assessments..[70] However, the Council believes that the Bulletin should explicitly compel, "to the greatest extent feasible, a formal uncertainty analysis [shall be done]." In support of this recommendation, the Council notes that when the National Research Council spoke directly to the topic of formal uncertainty analysis. They recommended that "…to the greatest extent feasible, EPA should present quantitative, as opposed to qualitative, representations of uncertainty."[71] [emphasis added]

The Council agrees with OMB that full risk characterization requires presentation of information on the general population (i.e., population 'at-large'). However, the Council finds the passage on page 19 to be confusing. The Bulletin implies that information regarding risk to the general population is required only "If highly exposed or sensitive subpopulations are highlighted…" The Council recommends that the passage be revised to read:

> "All assessments that are subject to this Bulletin must highlight risk estimates for the general population. If highly exposed or sensitive subpopulations are highlighted, the assessment must also highlight the general population to portray the range of variability."[72] [emphasis added to highlight changes]

The Council agrees with OMB that full risk characterization requires presentation of both "population risks" and "individual risks." However, a contradiction exists between two points that are directed to the topic of individual risk versus population risk.

[70] Bulletin at 25.
[71] Bulletin at 13 note 26.
[72] Bulletin at 19, with recommended changes noted.

> "…a risk manager may be interested in estimates of population and/or individual risk and an iterative dialogue would ideally bring this to the attention of a risk assessor early in the process."[73] [emphasis added]
>
> "When estimates of individual risk are developed, estimates of population risk should also be developed. Estimates of population risk are necessary to compare the overall costs and benefits of regulatory alternatives."[74] [emphasis added]

While the first suggests that a risk manager may declare an interest only in estimates of individual risk, the second passage clearly – and the Council believes, correctly – states that estimates of both individual and population risk are required for a high quality assessment. The contradiction can be resolved by revising the text from page 10 to read:

> "…while a risk manager may be interested in estimates of population and/or individual risk, both estimates are essential to full risk characterization (as specified in this Bulletin), and an iterative dialogue would ideally bring this to the attention of both risk assessor and risk manager early in the process."[75] [see emphasis added to highlight changes]

The draft Bulletin addresses the need for estimates of "countervailing risks and the Council concurs that full risk characterization requires this[76].

> "7. For risk assessments that will be used for regulatory analysis, the risk assessment also shall include:
> …
> b. a comparison of the baseline risk against the risk associated with the alternative mitigation measures being considered, and assess, to the greatest extent feasible, countervailing risks caused by alternative mitigation measures;
> …"[77] [emphasis added to highlight changes]

[73] Bulletin at 10.
[74] Bulletin at 16.
[75] Bulletin at 10, with recommended changes noted.
[76] Bulletin at 24, with recommended changes noted.
[77] Bulletin at 24.

APPENDIX 5

EXAMPLES OF RISK ASSESSMENTS THAT GROSSLY OVERSTATE RISK

Table of Contents

I. Human Health Risk Assessment Examples

A. Exposure Assessment

The first, and probably least controversial, task in human health environmental risk assessment (HHRA) is environmental sampling and analysis.[1] The second task is exposure assessment. This critical task involves measuring or estimating all parameters necessary to estimate, or model, mean and high-end human exposure to the chemical of interest through the scenario being assessed (e.g., exposure to a chemical that was disposed of in a landfill or exposure to combustion sources of air emissions). There are typically numerous steps in exposure assessment, including site/setting characterization, identification of fate and transport mechanisms, identification of potentially exposed populations, determination of direct and indirect (and complete and incomplete) exposure pathways, measurement/estimation of exposure parameter values or probability distributions (e.g., exposure frequency, exposure duration, or ingestion rate), and measurement or estimation of exposure point concentrations. Whenever possible, these steps should be based on site-specific data. Reliance on default assumptions should be avoided and used only when site-specific data are unavailable. In its Guidelines for Exposure Assessment, EPA (1992a) states:

> General default values should not be used in place of known, valid data that are more relevant to the assessment being done. The use of generic or surrogate data is common when site-specific data are not available. [However,] this is an additional source of uncertainty, and should be avoided if actual data can be obtained.

The final step in exposure assessment is calculating time-weighted intakes which are later combined with the toxicity assessment to characterize the risk. Time-weighted averages can be calculated using either a point estimate approach or a probabilistic method (such as Microexposure Event Modeling). For noncarcinogens, the time-weighted average is expressed as the average daily dose (ADD); for carcinogens, it is expressed as the lifetime average daily dose (LADD).

Although EPA is generally conservative in all aspects of exposure assessment, the Council has found that most of EPA's overconservativeness has been in estimating exposure parameters, including the magnitude, frequency and duration of exposure. The following examples illustrate the degree to which the Agency uses exposure factors that substantially overstate human exposure to chemicals in the environment.

[1] Steps in this task include collecting and compiling all existing, relevant data, determining whether additional data are required, collecting and analyzing new data, and data validation.

EPA Exposure Assessment Requirements in Connection with Petition for "Delisting" of Gas Turbines from MACT Requirements

Recently, the Gas Turbine Association petitioned EPA to "delist" gas turbines from MACT requirements pursuant to Section 112(c)(9) of the Clean Air Act on the grounds that gas turbine emissions present less than a 10^{-6} risk of cancer and an insignificant non-cancer risk. Section 112(c)(9) authorizes EPA to exempt source categories from MACT upon a showing that that no source in the category emits "hazardous air pollutants in quantities which may cause a lifetime risk of cancer greater than one in one million to the individual in the population who is most exposed to emissions of such pollutants" and, in the case of non-carcinogens, that emissions do not "exceed a level which is adequate to protect public health with an ample margin of safety and no adverse environmental impact will result." The petition is based on the fact that hazardous air pollutant (HAP) emissions from turbines are miniscule -- HAPs are emitted in the part per billion or part per trillion range, and the largest modeled impacts are less than 10% of ambient levels of HAPs.

GTA has been working to satisfy EPA's delisting requirements. Although EPA has stated that it is supportive of GTA's effort, the Agency is insisting that the risk assessment be conducted using exceedingly conservative exposure factors[2]:

- GTA's initial risk assessment used a 30-year exposure scenario. A 30-year exposure period is conventionally used in hazardous waste incinerator risk assessments to evaluate reasonable maximum exposure and is consistent with EPA's Risk Assessment Guidance for Superfund and Exposure Factors Handbook (EPA, 1989; EPA, 1997b). Based on national statistics, 30 years represents the upper-bound (95^{th} percentile) number of years that an individual might live at the same residence, while nine years represents the average number of years (50^{th} percentile). See Table 15-176 of EPA (1997b). EPA, however, has insisted that GTA use a lifetime exposure period of 70 years, with the first six years of exposure assumed to be childhood exposure.
- Although it is true that the statute requires that the delisting risk assessment be conducted using the most exposed individual, EPA is interpreting this requirement unreasonably. EPA insists that the risk assessment assume that the maximally exposed individual never leaves the point of maximum exposure (e.g., never goes to school or work) over a period of 70 years.
- EPA has demanded that GTA use the highest emission factor in EPA's database for turbine emissions, rather than the average.

[2] As discussed below, USEPA is also demanding that GTA use a very conservative approach to toxicity assessment.

Fox River Human Health Risk Assessment

The RI/FS for the Fox River/Green Bay Superfund Site was prepared by a contractor under oversight by the Wisconsin Department of Natural Resources (WDNR) and USEPA. The RI/FS includes a human health risk assessment (HHRA) of risks posed by PCBs in sediments at the site. The PCBs are taken up by fish and pass through the food chain to animals that feed on fish. Based on both the human health risk assessment and an assessment of ecological risks, USEPA and WDNR have issued a Record of Decision (ROD) calling for dredging of approximately five miles of the River. A ROD choosing a remedy for other portions of the River is expected this summer.

The HHRA substantially exaggerated human health risk in at least three respects. First, the HHRA relied on inappropriate estimates of the fish consumption rates of anglers who fish the Fox River. Second, the HHRA used unrealistic fish tissue concentrations because it: (a) relied on 1990's fish tissue data that do not reflect existing and future, declining, fish tissue concentrations; and (b) assumed, incorrectly, that the exposed population consumes a significant amount of carp. Third, the HHRA did not correctly adjust for the mobility of the population and thus overstated exposure duration. Based on these three factors, the HHRA likely exaggerates human health risk by several orders of magnitude. These factors are discussed briefly in turn.

Fish Consumption Rate

In order to assess risks to recreational anglers who consume fish from the Fox River fish, it is necessary to use fish consumption data that reflect the long-term consumption habits of recreational anglers who actually use the fishery. However, the studies relied on by the HHRA (West et al. 1989a; 1993) do not provide data for people who fish the Fox River and do not provide reliable estimates of long-term consumption rates because the studies collected only short-term data. The study design used in the West et al. studies collected data on the consumption habits of Michigan anglers using a one-week recall period. As EPA acknowledges in its Exposure Factors Handbook (EPA, 1997b), "the distribution of average daily intake reflective of long-term consumption patterns cannot in general be estimated using short-term (e.g., one week) data." EPA (1997b) itself has concluded that the West et al. (1993) study should not be used to estimate long-term consumption rates, stating that "the resulting distribution [of the West et al. (1993) study] will not be indicative of the long-term fish consumption distribution and the upper percentiles reported from the EPA analysis will likely considerably overestimate the corresponding long term percentiles" (EPA, 1997b).

Using upper percentile consumption rates from the West et al. (1993) data resulted in significantly inflated consumption rates. The HHRA should instead have used readily-available data from the Wisconsin Fishing and Outdoor Recreation Survey (WFORS) which provides long-term consumption data obtained from Wisconsin, rather than Michigan, anglers. In this study, data were recorded by anglers in diaries that were maintained over a period of four months. This data collection methodology both minimized potential recall bias and provided data on long-term behavior. As a result, far fewer simplifying assumptions are needed to extrapolate

these data to annualized fish consumption rates. Using the WFORS data rather than the West at al. data would lower the Fox River risk estimates by at least a factor of two.

PCB Concentration in Fish

The HHRA used inappropriately high estimates of PCB fish concentrations because: (a) it relied on 1990's fish tissue data that do not reflect existing or future, declining concentrations of PCBs in Fox River fish; and (b) it assumed, incorrectly, that the angler population consumes a significant amount of carp, a fish that accumulates higher concentrations of PCBs than other species. These errors increased the HHRA's estimates of risk by the following factors.

- Failure to use current and future PCB concentrations in fish – factor of 10. The HHRA averaged PCB fish concentrations for the 1990s, and assumed that these concentration remained static into the future. Large-scale fish sampling conducted in 1998 by the Fox River Group (FRG) demonstrated significant declines in fish tissue PCB concentrations. Using principally these 1998 data for purposes of calculating risk resulted in a 10-fold reduction in PCB fish tissue concentrations, on average. Moreover, results of food web modeling suggest that PCB tissue concentrations will continue to decline over time.
- Assuming high consumption of carp – factor of 1.3 to 30. The HHRA assumed that recreational anglers consume a large amount of carp. The FRG's review of angler surveys for the Lower Fox River indicated that carp were rarely caught and eaten. Because carp contain a higher relative percentage of lipids, they generally also have higher PCB tissue concentrations.

The total impact of these two factors is multiplicative, so the total increase in risk estimates based on errors in defining PCB concentration in fish is between 13 and 300.

Population Mobility

The HHRA assumed exposure durations that were far too long because it ignored population mobility. For the hypothetically highly-exposed individual (referred to as the "Reasonable Maximum Exposed" individual or "RME" individual) and the typically exposed individual (referred to as the "Central Tendency Exposure" individual or "CTE" individual) the HHRA assumed exposure durations of 50 and 30 years, respectively. There is no valid basis for these exposure durations. USEPA's Exposure Factors Handbook (EPA, 1997b) recommends point estimate exposure duration values for the RME and CTE individuals of 30 and 9 years, respectively, and notes that at least three studies support these values. The HHRA's departure from the accepted exposure values results in increases of the RME risk by a factor of more than 1.5 and increases of the CTE risk by a factor of more than 3.

The combination of these errors results in exaggeration of the risk posed by consumption of Fox River fish by as much as two orders of magnitude.

Upper Hudson River Human Health Risk Assessment[3]

EPA's HHRA for the Upper Hudson River substantially overestimates human health risk for several reasons. Primary among these is that the HHRA grossly overstates the rate at which Upper Hudson River anglers might consume fish from the River.

Because EPA failed to conduct a fish consumption survey of anglers who might use the Hudson River, it was forced to rely on studies of other water bodies and angler populations. However, although there are five studies that might have been used to estimate Upper Hudson River angler fish consumption, EPA chose to rely on a single study, Connelly et al. (1992). This study reported rates of fish consumption that are about four times higher than the average of the other studies (EPA, 2000f).

EPA's use of the 1992 Connelly study to estimate fish consumption rates was inappropriate for several reasons. First, EPA derived a consumption rate for fishermen almost three times greater than the authors of the study found. The paper states that the average number of meals consumed by responding anglers was 11 meals per year which, using a 0.5 pound meal size, results in a mean consumption rate of 6.8 g/day instead of the 17.3 g/day calculated by EPA (Connelly, et al., 1992). Next, the study was not designed to assess consumption rates, but rather angler awareness of and knowledge about fish consumption advisories. As a result, numerous assumptions were required to generate consumption rates (e.g., meal size, types of Hudson River fish eaten, the type of waterbody the surveyed anglers fished in, etc.). Third, individuals who do not respond to surveys of this type are likely to consume considerably less fish than individuals who do respond (Connelly et al., 1992; West et al., 1989a,b). The 52.3% response rate reported by Connelly is on the low-end of acceptable standards, which biases fish consumption estimates toward higher level consumers, leading to an overestimate of fish consumption rates. EPA itself recognized that some of the rates generated in the 1992 Connelly study were beyond credibility. The Agency discarded some of the high end consumption results and used the 90th percentile, rather than the usual 95th percentile, in its point estimate for the high end of exposure (EPA, 2000f). These and other limitations led EPA itself to conclude that the study should not be considered a "key" study when evaluating freshwater fish consumption by recreational anglers (EPA, 1997b).

Most importantly, consumption rates based on Connelly et al. (1992) are inconsistent with well-conducted studies of similar angler populations which are more appropriate for estimating rates of fish consumption for the Upper Hudson. See Table 1.

[3] As set forth in EPA's "Phase 2 Report – Review Copy, Further Characterization and Analysis, Volume 2f – Human Health Risk Assessment, Hudson River PCB Reassessment RI/FS."

Table 1: Comparison of Fish Ingestion Rates (g/day) from Studies of Northeastern Recreational Anglers

Consumption Rate Percentile	Connelly et al. (1992) New York Multiple Rivers[a]	Ebert et al. (1993) Maine Multiple Rivers	ChemRisk (1991) Maine Single River[b]	Connelly et al. (1996) New York All Waters[c]	Ebert et al. (1996) Connecticut Single River[d]
50th	4.0	0.99	0.49	2.2	0.17
90th	31.9	6.1	5.3	13.2	5.8
95th	63.4	12.4	10.7	17.9	12
Arith. Mean	17.3	3.7	3.0	4.9	2.6

a. EPA (2000f) analysis
b. West Branch Penobscot River
c. EPA (2000f) analysis
d. Housatonic River

Each of the studies listed in the table (other than Connelly et al., 1992) was designed specifically to assess rates of fish consumption. Further, for example, in their 1996 study, Connelly and co-workers substantially reduced the possibility of recall bias by using food diaries, which tend to better represent long-term consumption habits (Connelly et al., 1996). Moreover, Connelly et al. (1996) and Ebert et al. (1996) also had higher rates of response and were, therefore, more representative of the targeted angler population. Finally, with any of the four alternative studies, there is no need to assume an arbitrary meal size in order to derive consumption estimates. Based on these facts, it is clear that using Connelly et al. (1992) skewed EPA's estimates of exposure. The evidence of the relevant angler surveys taken as a whole is that realistic estimates of fish consumption are approximately one-quarter of those assumed by EPA: i.e., 1 meal per month for the high exposure angler.

Housatonic River -- Upper Two Miles Human Health Risk Assessment

In May 1998, EPA and the Commonwealth of Massachusetts Department of Environmental Protection (MDEP) published a memorandum titled "Evaluation of Human Health Risks from Exposure to Elevated Levels of PCBs in Housatonic River Sediment, Bank Soils and Floodplain Soils in Reaches 3-1 to 44-6 (Newell Street to the Confluence of the East and West Branches)" (Housatonic Two-Mile HHRA). This HHRA covered a two-mile stretch of the Housatonic River within the City of Pittsfield, Massachusetts, and was used by EPA as the basis for requiring cleanup of that stretch of the river. The HHRA envisioned three hypothetical receptor scenarios involving exposure to PCBs along different areas of the river reach in question:

- EPA assumed that in Exposure Area A, where the river is bordered primarily by commercial properties, 9 to 18-year-old children would trespass along the river banks ("youth trespasser scenario").
- EPA assumed that in Exposure Area B, where the river is bordered by residential properties and the banks are moderately steep, children between the ages of 5 and 12 years of age would wade in the water and play in and along the river banks ("child wader scenario").
- EPA assumed that in Exposure Area C, where the river is bordered by residential properties and the banks are not steep, children between the ages of 1 and 6 years old would wade in the water and play in and along the river banks ("child resident scenario").

EPA then estimated cancer and non-cancer risks to these hypothetical populations using a variety of exposure assumptions. Among these assumptions were that:

- The youth trespasser would contact soils and sediments along the river two days per week every week from April through October (61 days per year) for nine years. Each time he visited the river, he would get soil or sediment all over his hands, arms, feet and lower legs (843 square inches – or almost 6 square feet -- of skin). The soil/sediment would remain on the skin for 24 hours.[4] The youth would also eat 50 milligrams of soil each and every day he visited the river. Finally, the soil that was contacted or ingested would always be contaminated with virtually the highest concentration of PCBs that had been detected in soils and sediments of the river reach (rather than the average).[5]

[4] This assumption is inherent in the agencies' use of the assumption that 14% of the PCBs contained in the soil which the children contact would be absorbed through the skin. This factor was taken from Wester et al. (1993), in which monkeys were found to absorb this fraction of PCBs over a period of 24 hours when PCB-containing sand was stuck to a shaved area of their bodies with a patch. (Note that this sand had a very low organic carbon content, which would tend to make PCBs more bioavailable than they would be in soil with higher organic carbon; a more recent study of monkeys using soil that contained an organic carbon content more typical of U.S. soils found that the absorption rate was only approximately 4%. Mayes et al. (2002).)

[5] For exposure point concentration, the agencies used the 95 percent upper confidence limit ("UCL_{95}") on the mean, unless this value exceeded the maximum concentration, in which case the maximum was used. These values were much higher than the actual mean. For example, for Exposure Area A, the mean sediment and soil concentrations were 17 and 275 ppm, respectively. The values used by the agencies were 46 and 2,400 ppm, respectively.

- The child wader was assumed to contact soils along the river five days per week every week from April through October (153 days per year) and to contact sediments in the river five days per week every week from June through August (65 days per year). This remarkably consistent behavior would go on for seven years. Like the youth trespasser, the child wader would get soil or sediment all over a large portion of her body (570 in^2 of skin for soil and 832 in^2 of skin for sediment) each time she visited the river, the soil/sediment would adhere for 24 hours, she would eat 50 milligrams of soil/sediment per day of exposure, and the soil/sediment contacted or ingested would be the most contaminated soil or sediment available.[6]
- The same exposure frequency assumptions that were used for the child wader were used for the child resident, and the behavior was assumed to last five years. The child resident was assumed to get even dirtier than the child wader in proportion to her size, having 445 in^2 of skin exposed to soil and 661 in^2 of skin exposed to sediment for 24 hours. The child resident would eat 100 milligrams of the most contaminated soil or sediment available on each such day.[7]

These exposure assumptions are clearly excessive in terms of PCB concentration, exposure frequency[8], exposure duration, and extent of skin exposure. But just as important, as explained in a report titled "Critique of Agencies' Human Health Risk Assessment for the Two-Mile Reach (July 7, 1998)" (Critique) that was prepared by ChemRisk and submitted to EPA and MDEP by General Electric, actual blood PCB concentration data taken from residents of Pittsfield, including those who lived along the river, showed that these residents who did not have occupational exposure to PCBs did not have elevated levels of PCBs in their bodies. Rather, their blood PCB concentrations were within the background range (mean of 4 to 8 ppb) for non-occupationally exposed populations in the U.S.[9] Thus, the agencies' human health risk assessment for this reach of the Housatonic River was not only overconservative, but also inconsistent with the actual empirical data.

[6] For example, for Exposure Area B, the mean sediment and soil concentrations were 89 and 36 ppm, respectively. The values used by the agencies were 905 and 377 ppm, respectively. Using the 905 ppm value is particularly egregious. Four samples taken in very close proximity to the location where the 905 ppm concentration was found had PCB concentrations ranging from 1.6 to 51 ppm, or 18 to 566 times lower.

[7] For example, for Exposure Area C, the mean sediment and soil concentrations were 16 and 23 ppm, respectively. The values used by the agencies were 30 and 68 ppm, respectively.

[8] For example, Exposure Area A is primarily commercial, with riverbanks that are steep, heavily vegetated, fenced and posted with signs that warn of the presence of PCBs. The HHRA assumed that individuals would frequent this area two days per week for seven months of the year. Although individuals might conceivably trespass here occasionally, it was unreasonable to assume that individuals would trespass on such a regular basis.

[9] These findings are consistent with studies that have been performed near other PCB-contaminated sites in Norwood, Canton, and Fairhaven Massachusetts, Paoli, Pennsylvania, Milford, New Hampshire, and Bloomington, Indiana.

Manistique Harbor Contaminated Sediment Site

The Baseline Human Health Risk Assessment (BHHRA) for the Manistique Harbor Sitecalculated a cancer risk to the average and "high-end" recreational angler of 1.8 x 10-5 and 2.4 x 10-3, respectively. {Note: Using the current CSF for PCBs of (2.0 mg/kg/d)-1, these cancer risks would be 4.6 x 10-6 and 6.2 x 10-4, respectively]. For the "average" and "high-end" subsistence anglers the risks were 2 x 10-4 and 1.2 x 10-2, respectively. These cancer risks estimates were used by Region 5 risk managers to require remediation of the Manistique Harbor sediments.

The BHHRA incorporated numerous overly conservative assumptions and the result was a remedy that has to date cost in excess of $48 million. If a more reasonable, appropriate, and scientific assessment of the Harbor was conducted, it would have been determined that the levels of PCBs in surface sediments posed significantly less or no risk to human health or the environment. This would likely have led to a less rigorous remedy, or possibly a determination that remediation was not needed.

The folloing are examples of some of the redundant conservatism inherent in the BHHRA:

- The high-end angler scenarios (both recreational and subsistence) assumed that 25% of the diet (consumption of 54 and 130 g/day every day for 30 years) was carp, despite the finding that few if any Upper Peninsula anglers regularly consume carp (West et al., 1993) . The impact of this unfounded assumption was significant since the available fish tissue sampling data showed that carp consistently contained the highest concentrations of PCBs of all fish species sampled in Michigan. For example, the tissue concentration for carp used in the Manistique BHHRA was 6.5 mg/kg; but the walleye tissue concentration was only 0.34 mg/kg.
- It was assumed that subsistence anglers obtained 50% (average exposure scenario) or 100% (high-end scenario) of their fish from Manistique Harbor. This was unlikely given the demographics of the population and the difficulty associated with fishing from the banks of the Harbor. Manistique Harbor is small, the banks are bulkheaded, and better and more accessible fishing areas on Lake Michigan are readily available. It was also assumed that the anglers consumed fish from the Harbor 365 days a year. Since the Harbor freezes over in the winter, this assumes that a substantial amount of fish is caught, saved, and consumed over the winter. An informal survey of the resident community could not identify any individuals engaged in this level of fishing and consuming fish from the Harbor.

Consideration of more reasonable and scientifically-based exposure assumptions and toxicity factors would have demonstrated that the surface sediment concentrations of PCBs in the Harbor did not represent a significant risk to the local populations. One must question the benefit to society (locally and nationally) of remediating a harbor to PCB fish tissue levels of less than 0.5 mg/kg (as in the walleye) and average surface sediment concentrations basin wide (56 acres) of 5.2 mg/kg at the cost of over $48 million.

EPA's Recent Increase in the Default Assumption for Adult Soil Ingestion Rate

EPA recently increased its default assumption for adult soil ingestion rate from 50 mg/day to 100 mg/day (EPA, 2001c; EPA; 2002b). As justification for this change, the Agency cites its Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual, Supplemental Guidance: Standard Default Exposure Factors (EPA,1991). This guidance relies on Calabrese et al. (1990) for the 100 mg/day figure. EPA's Exposure Factors Handbook (EPA, 1997b), on the other hand, continues to recommend an adult soil ingestion rate of 50 mg/day. The Exposure Factors Handbook value is also based primarily on Calabrese et al. (1990).

There is no basis for EPA's increase in the adult default soil ingestion rate or for the Agency's changing its mind on how to interpret Calabrese et al.(1990). The Exposure Factors Handbook accurately cites Calabrese et al. (1990) as having found a soil ingestion range of 30 to 100 mg/day and supporting a "reasonable central estimate of adult soil ingestion" of 50 mg/day. Be that as it may, the most recent work on this subject by Calabrese et al. (1997) and Stanek et al. (1997) supports an adult soil ingestion range of 20 to 40 mg/day. The authors of Stanek et al.(1997) also highlight many significant improvements over their previous work (i.e., Calabrese et al., 1990), supporting the use of the 1997 ingestion rates for risk assessment purposes. Thus, there is no scientific basis for use of a default soil consumption rate of 100 mg/day.

Refusal to Alter Construction Worker Soil Ingestion Rate

Historically, EPA and state regulatory agencies have used a soil consumption rate of 480 mg/day in assessing construction worker exposure to soil contaminants through the soil ingestion exposure route. Although EPA Headquarters has determined that this factor is excessive, Headquarters has to date not published guidance encouraging the EPA Regions and the states to use a more realistic value.

The soil ingestion value of 480 mg/day is based on Hawley (1985). The Hawley (1985) estimates of soil consumption rates are based not on empirical data, but rather on various assumptions about the extent to which soil/dust adheres to hands, hand-to-mouth behavior, and frequencies of indoor and outdoor activities. Hawley (1985) stated that his estimates of soil ingestion were subject to substantial uncertainty. EPA has stated that the Hawley (1985) estimates must be considered "conjectural" due to the lack of supporting measurements (EPA, 1997b).

To derive soil ingestion rate estimates, Hawley (1985) assumed that an adult, while engaged in yard work or other physical activity such as construction excavation, would incidentally ingest half of the soil that could coat the inside surfaces of his fingers and thumbs on both hands twice per day. He further assumed that soil would adhere to the skin at a rate of 3.5 milligrams of soil per square centimeter of skin surface (mg/cm^2). This estimated loading or adherence factor was based an approximation of both the density of the soil and the thickness of a layer of soil that might adhere to the hands of an individual in contact with the ground.

Sheppard (1995) showed that the Hawley (1985) estimate were excessive. Sheppard (1995) demonstrated that the soil loading assumption of 3.5 mg/cm^2 for the arms and hands would result in a very high and conspicuous soil load. Sheppard reported that a load of less than 1 mg/cm^2 was more reasonable because a load greater than 1 mg/cm^2 would be highly noticeable and would deter hand-to-mouth contact. Sheppard (1995) also noted that Hawley (1985) did not account for the fact that individuals with substantial amounts of soil on their hands would be unlikely to pick up food or put their hands in their mouths before washing or wiping their hands.

Recently, new data on soil adherence have been published. As reported by EPA (1997b), Kissel et al. (1996) and Holmes et al. (1996) directly measured the amount of soil that adheres to skin surfaces during a variety of occupational and recreational activities. These studies indicated that the amount of soil that adheres to the skin depends on the type of activity performed and the body parts that come into contact with the soil. As one would expect, soil adherence to the skin appears to be greatest during outdoor activities such as farming and gardening, and more soil/dust tends to adhere to the hands than to other areas of the body.

Using the data provided by Holmes et al. (1996), EPA's dermal workgroup derived average (geometric mean) and high-end (95th percentile) adherence rates of 0.24 mg/cm^2 and 0.468 mg/cm^2, respectively, for the hands of construction workers and recommended that these values be used in risk assessment (EPA, 1997b; EPA, 2001f). Even the high-end adherence factor

(0.468 mg/cm^2) measured by Holmes et al (1996) is considerably lower than Hawley's estimate of 3.5 mg/cm^2.

If one uses EPA's (1999c; 2001f) average (0.24 mg/cm^2) and high-end (0.468 mg/cm^2) soil adherence values for the hands of construction workers in place of Hawley's assumed adherence value of 3.5 mg/cm^2, but retains all of Hawley's other exposure assumptions, the resulting construction worker soil ingestion rate ranges from 33 mg/day (average) to 64 mg/day (high-end). Thus, a conservative soil ingestion rate for a construction worker is 64 mg of soil per day, not the value of 480 mg/day which is still embraced by the EPA regions.

Continued Use of Excessive Dermal Absorption Rate for PCBs

In evaluating dermal absorption of a chemical in soil, EPA selects a dermal absorption factor to estimate the fraction of the chemical in the soil adhering to the skin that will actually be absorbed. For PCBs, EPA (2001f) assumes a 14 percent absorption factor based on a study by Wester et al. (1993). This absorption factor likely overestimates the fraction of PCBs that are absorbed due to the following limitations with the study.

- The Wester et al. study evaluated dermal uptake from larger grained sandy soil that had an organic carbon content much lower than that typically found in soils and sediments at PCB-contaminated sites. As demonstrated by Roy et al. (1990), higher levels of organic carbon in soils substantially decrease the bioavailability of PCBs for dermal absorption.
- The study methodology did not mimic chemical mixtures or conditions of dermal exposure that would be expected to occur during and after actual exposures.
- A number of studies have shown that a significant fraction of PCBs that have been in soil or sediment for a considerable period of time become tightly bound to the soil or sediment and desorb quite slowly, thus reducing their bioavailability. Wester et al., however, evaluated absorption of freshly spiked PCBs, an approach that does not emulate the fate of aged PCBs in site soils.
- Studies have shown that dermal uptake by monkeys is greater than uptake through human skin and that permeability of abdominal skin (as was tested by Wester et al.) is much greater than permeability of the extremities – the skin areas most likely to be in contact with site soils and sediments.

These limitations indicate that the 14 percent dermal absorption factor derived from Wester et al. (1993) likely overestimates the degree of dermal absorption of PCBs. General Electric sponsored a study to evaluate several aspects of the Wester et al. (1993) study in order to estimate PCB dermal absorption for application in the risk assessment for the Housatonic River site. This study was conducted by Huntingdon Life Sciences using Rhesus monkeys as test animals and soil taken from the Housatonic River floodplain that was spiked with Aroclor 1260, the PCB mixture most common to the Housatonic River site. The floodplain soil had an organic carbon content of 5-6 percent, which is typical of most soil and is in contrast to the 0.9 percent organic carbon content in the soil used by Wester et al. Because Wester et al. used soil that had been freshly spiked with PCBs, the Huntington study evaluated the relative rate of PCB absorption from freshly spiked PCB soil versus that from PCB-containing soil that had been aged to simulate weathered PCB soil. Finally, since the 24-hour exposure period evaluated by Wester et al. seems likely to have overestimated the period that humans would be dermally exposed to PCB-containing soil before washing, the Huntington study included a 12-hour dermal exposure period as well as 24-hour exposures. All other aspects of the Huntington study were similar to those used by Wester et al.

Using the same procedure as Wester et al. to calculate dermal absorption rates, the calculated mean dermal absorption rates for PCBs from soil in the Huntington study were:

Group exposed for 12 hours to aged PCBs in soil: 3.43 percent
Group exposed for 24 hours to freshly spiked PCBs in soil: 4.07 percent
Group exposed for 24 hours to aged PCBs in soil: 4.26 percent

As the results show, the use of the default 14 percent dermal absorption factor is not appropriate for a risk assessment of the Housatonic River; rather a dermal absorption factor of approximately 4 percent would be recommended based on site-specific data. Moreover, 4 percent would be a more appropriate factor at any site where the organic carbon content of the soil is similar to that in the Huntingdon study.

To date, EPA has not accepted the dermal absorption factor of 4 percent and still relies on its default value of 14 percent.

EPA Failure to Perform Probabilistic Risk Assessments

The evaluation of variability and uncertainty is an important component of the risk characterization task of risk assessment. As stated in the 1995 Risk Characterization memorandum from Administrator Carol Browner (EPA, 1995b):

> [W]e must fully, openly, and clearly characterize risks. In doing so, we will disclose the scientific analyses, uncertainties, assumptions, and science policies which underlie our decisions . . . . There is value in sharing with others the complexities and challenges we face in making decisions in the face of uncertainty.

EPA's Risk Assessment Guidance for Superfund (RAGS) Volume III: Part A (EPA, 2001a) provides technical guidance on the application of probabilistic risk assessment (PRA) methods to human health and ecological risk assessments in the Superfund program. The guidance focuses on Monte Carlo analysis (MCA) as a method of quantifying variability and uncertainty in risk. In addition, the 1997 EPA Policy for Use of Probabilistic Analysis in Risk Assessment (EPA, 1997c) states:

> It is the policy of the U.S. Environmental Protection Agency that such probabilistic analysis techniques as Monte Carlo analysis, given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments.

According to EPA (2001a), the decision to use probabilistic analysis is

> site-specific and is based on the complexity of the problems at the site, the quality and extent of site-specific data, and the likely utility of the result. If the additional information provided from a PRA is unlikely to affect the risk management decision, then it may not be prudent to proceed with a PRA. However, if there is a clear value added from performing a PRA, then the use of PRA as a risk assessment tool generally should be considered despite the additional resources that may be needed.

EPA (2001a) endorses advanced modeling approaches for characterizing variability and uncertainty. According to EPA (2001a), an example of a more advanced approach is Microexposure Event Analysis (MEE):

> Where information is available to characterize variability on a smaller time scale than life-time, an alternative expression of dose that accommodates such variability may be desirable. Daily activity patterns, food intake, soil ingestion and other behavioral factors are measured in a time period of less than a year. The extrapolation of these short term results to the chronic exposure situation is a source of uncertainty. Exposure events are real but

> unknowable, whereas data regarding the nature and magnitude of these events is known but its application to a real world situation is uncertain. Microexposure event analysis (MEE) attempts to explicitly quantify this uncertainty. MEE modeling provides an alternative to the standard time-averaging approach. In the MEE approach, long term intake is viewed as the sum of individual exposure events. Implementing the MEE approach in a PRA requires dividing the exposure duration into short epochs, or time steps, within which the values assigned to exposure variables remain constant, but are allowed to vary from one time step to the next.

In other words, MEE captures temporal changes in inter-individual variation (Simon, 1999).

The theory and methodology of the MEE model are documented in the literature (Harrington et al., 1995; Price et al., 1996; Keenan et al., 1996; Simon, 1999). In brief, an individual's total exposure to a contaminant is calculated by summing the doses received during many individual exposure events. Each individual event is simulated using information specific to the time and location of the exposure events. The number of events and sequence in which they occur in the person's life can be simulated based upon information about an individual's short- and long-term behavior. This approach avoids the difficulty inherent in conventional Monte Carlo modeling where unrealistic exposures may be projected due to coincidental selection of the upper percentiles of two or more input distributions.

Although EPA's guidance endorses PRA, including Monte Carlo analysis and MEE, EPA often EPA views probabilistic analysis as a means to support point estimate risks, rather than as a way to more accurately characterize risks. As examples, two prominent Superfund sites involving sizeable river systems with PCB-contaminated sediments – the Upper Hudson and Fox Rivers -- are discussed below.

Upper Hudson River Probabilistic HHRA

EPA's probabilistic model for the Upper Hudson River included deficiencies in both the model design and documentation of the assessment. The design of the HHRA model forced EPA to assume that anglers consumed unrealistic amounts of fish harvested from the same locations, cooked in the same fashion, and composed of the same mixture of species every year for more than 30 years. The model did not account for the way in which people's behavior would vary over time, nor did the HHRA account for declining concentrations in fish tissue PCB levels in the future.

Had EPA followed its own guidance and used an MEE model, it would have generated more realistic estimates of exposure from ingesting Upper Hudson River fish. Exposures to Hudson River anglers should have been modeled as a series of separate exposure events that occur over time, taking into consideration temporal changes in fish tissue concentrations and angler behaviors. Furthermore, had EPA used the MEE model to compare the benefits of several remedial alternatives at reducing risks to the hypothetical fish consuming angler, it would have demonstrated that dredging produces no additional risk reduction compared to source control and that source control achieves acceptable PCB concentrations in fish for the average Upper Hudson River angler ten years sooner than does dredging in 29 of the 40 miles of the Site. Moreover, the MEE model demonstrated that source control achieves lower risks to human health (both cancer and non-cancer) than dredging in 34 of the 40 miles of the Site. EPA ignored the results of this analysis and issued a ROD for the Upper Hudson River which includes a dredging project of unprecedented scale which is projected to cost well over $500,000,000.

Fox River Probabilistic HHRA

EPA Region 5 and Wisconsin Department of Natural Resources (WDNR) conducted a HHRA that relied principally on a point estimate or deterministic approach in arriving at estimates of cancer and noncancer risk from consuming fish from the Lower Fox River. Although the EPA/WDNR HHRA used certain probabilistic methods as part of a sensitivity analysis, the HHRA did not include a true probabilistic risk assessment. As noted earlier in these comments, many of the input assumptions used in the EPA/WDNR HHRA, especially those relating to fish consumption rates, PCB fish tissue concentrations, and population mobility, were flawed or based on outdated information. Consequently, due to the multiplicative nature of deterministic assessments, the HHRA results overestimated risk by up to several orders of magnitude.

The Fox River Group companies (FRG) prepared an alternative Human Health Risk Assessment of the Lower Fox River and Green Bay (AMEC, 2002a). The starting point for the FRG's HHRA was the use of reliable scientific data from Fox River fish, sediments, and water that reflect current conditions and historical trends. The risk assessment incorporated a large media sampling database, Wisconsin fish consumption data, age- and region-specific data on human mobility, and state-of-the-art fate and transport, food web, and risk assessment models. In accordance with EPA (2001a) guidance, the FRG conducted an advanced Microexposure Event (MEE) probabilistic risk assessment because such an analysis adds value whenever screening risk estimates are above levels of concern and when the costs of remediation are high. Output from the MEE model showed that estimated risks were lower than the risks calculated in the EPA/WDNR HHRA by at least an order of magnitude, and in some cases, by as much as two orders of magnitude. Furthermore, when the MEE model was used to compare various remedial alternatives, it demonstrated that the proposed massive dredging remedy would offer no measurable benefit at reducing human health risks to anglers who fish the Fox River or Green Bay.

B. Toxicity Assessment

Toxicity assessment begins with a review of all relevant studies regarding the toxicity of the chemical at issue through all relevant routes of exposure (e.g., oral, dermal or inhalation). The next step is determining the critical effect – generally, the health effect caused by the chemical which occurs at a dose lower than the doses that cause any other health effects that the chemical may have. Then, a decision is made regarding the study data which will be used to quantify the toxicity of the chemical. If the chemical is a noncarcinogen, a NOAEL, LOAEL or Benchmark Dose Lower Limit (BMDL) is determined from the study data. The LOAEL, NOAEL or BMDL is, if necessary, converted from a measure of concentration (e.g., µg/L of blood) to a measure of dose (e.g., µg/kg body weight/day). Finally, the LOAEL, NOAEL or BMDL is divided by one or more uncertainty factors to yield the RfD, the ultimate product of the toxicity assessment.

If the chemical is a carcinogen, a model is used to derive a cancer slope factor (CSF) from the tumor count data from the chosen study. Essentially, a so-called "best-fit" line is drawn between the responses observed in an animal bioassay and that line is then extrapolated under a linear low-dose response assumption in order to predict the dose that would be anticipated to produce either a 10 percent (ED_{10}) or a 1 percent (ED_{01}) response rate in the population of test animals. Next, the linear low-dose response model is used to predict the statistical 95 percent lower confidence bound of the ED_{10} or ED_{01} (the LED_{10} or LED_{01}, respectively). Finally, a straight line, linear low-dose extrapolation is performed between either the LED_{10} or LED_{01} and zero, and the slope of this line is the cancer slope factor or CSF.

Toxicity assessments are typically not performed by EPA in the course of risk assessments when EPA's Integrated Risk Information System database (IRIS, 2003) contains a current RfD or CSF for the chemical at issue. However, EPA risk assessors have been advised to consider all available data in the course of risk assessments as well as to perform a toxicity assessment when new data are available that were not originally considered in developing the IRIS RfD or CSF (EPA, 1993).

EPA toxicity assessments are often over-conservative for several reasons. For carcinogens, a pervasive problem has been use of a linear multi-stage (LMS) model to estimate low-dose cancer risk from high-dose animal studies. Although EPA guidance does not require use of this model where dose-response and/or mechanistic information exist, the Agency has often been unwilling in practice to depart from use of the LMS model. In the case of non-carcinogens, the most prevalent problem is the magnitude and number of uncertainty factors used to derive an RfD from study dose-response data. Another problem arises from the Agency's refusal to abandon misconceptions regarding the hazards posed by a chemical even in the presence of long-term human data showing that the chemical does not cause adverse health effects at doses that EPA predicts to be harmful. Some examples of these problems, all of which result in overestimates of chemical toxicity, follow.

EPA Cancer Risk Assessment Procedures

The Executive Summary to these comments quotes an EPA guidance document for the proposition that EPA risk assessments typically overpredict risk:

> To account for these uncertainties and to acknowledge gaps in science, we build in safety factors in the risk estimates which tend to overestimate what we believe to be the actual risk. Where there is uncertainty or where our information is incomplete, we make assumptions that tend to overestimate the risks as a way to insure the public health is protected.

EPA (2000a). Nowhere is this more likely than in the area of cancer risk assessment. EPA's 1986 cancer risk guidelines states that:

> It should be emphasized that the linearized multistage procedure leads to a plausible upper limit to the risk that is consistent with some proposed mechanisms of carcinogenesis. Such an estimate, however, does not necessarily give a realistic prediction of the risk. The true value of the risk is unknown, and may be as low as zero.

EPA (1986a).

Over the last several years, EPA has been in the process of revising its cancer risk assessment guidelines. Although the revisions, including the 2003 draft final guidelines (EPA, 2003a), have sought to bring more precision to the process of assessing the human cancer risk of chemicals, the processes set forth in the guidelines continue to inject a high degree of conservatism into carcinogenicity assessment. In fact, although EPA continues to admit that its carcinogen risk assessment procedures, particularly use of the linearized multistage (LMS) model, may predict risk when none exist, EPA is actually playing-down the problem. There are so many conservative aspects of EPA's cancer risk assessment procedures that EPA cancer assessments typically predict substantial risk when none exists. The following provides a brief summary of seven of EPA's cancer assessment procedures that combine to grossly exaggerate cancer risk from exposure to chemicals:

- Although the use of the LMS as a default, in itself, tends to overstate risk in virtually all cases, EPA compounds the problem by stating its preference for a default "Lower Limit on Effective Dose" (LED10), defined as the "lower 95% limit on a dose that is estimated to cause a 10% response" (EPA, 2003a). According to EPA, the LED10 is a "protective" level to account for experimental variability. Although the LED 10 is certainly protective, the EPA Science Advisory Board (SAB) (SAB, 1997) and the American Industrial Health Council (American Industrial Health Council, 1999) have recognized that using the LED 10 as the departure point in carcinogen risk assessment injects a layer of undue conservatism in what should be a scientific, not a public policy, exercise. Accordingly, these groups have urged EPA to use the central estimate "Effective Dose (ED)" (ED10) as the point of departure in an effort to most accurately characterize risk. EPA disagrees, arguing that the

ED10 is appropriate only for use in ranking the relative hazard/potency of agents for priority setting. The Council submits that this position is driven by nothing other than an effort to inject conservatism into the risk assessment process instead of leaving public policy judgments to the risk management phase.

- EPA guidance allows departure from the LMS when sufficient mode of action data are available: "When adequate data on [mode of action] show that linearity is not plausible, and provide sufficient evidence to support a nonlinear [mode of action] for the general population and any subpopulations of concern, the default changes to a different approach – a reference dose/reference concentration – that assumes that nonlinearity is more reasonable" (EPA, 2003a). Note that EPA's statement contains a significant qualifier that will very rarely be met – the LMS may be abandoned only when a nonlinear mode of action is supported for both the general population and any "subpopulations" that EPA may deem to be "of concern." Moreover, EPA again insists that the point of departure be the conservative LED 10, not the central tendency ED 10 (EPA, 2003a). Finally, EPA states that "the point of departure mostly will be from . . . precursor response data, for example hormone levels of mitogenic effects rather than tumor incidence data" (EPA, 2003a). This means that EPA will predict carcinogenicity from possible precancer effects, rather than from actual tumor data. Thus, EPA will allow risk assessors to depart from the default assumption of the ultraconservative LMS – but only if they then inject additional levels of conservatism into their risk assessment.

- In its treatment of epidemiological data, EPA states that its default position is as follows: "When cancer effects in exposed humans are attributed to exposure to an exogenous agent, the default assumption is that such data are predictive of cancer in any other exposed human population" (EPA, 2003a). This assumption seems fair. But EPA also states: "When cancer effects are not found in an exposed human population, this information by itself is not generally sufficient to conclude that the agent poses no carcinogenic hazard to this or other populations of potentially exposed humans, including susceptible subpopulations or life stages" (EPA, 2003a) (emphasis added). In part, this statement also seems fair – for example, it would not necessarily be appropriate to conclude from a negative cancer occupational exposure study that the substance in question does not pose carcinogenic risks to infants.[10] But EPA is saying more than that – the quoted language states that when cancer effects are not found in an exposed human population, this information may not be sufficient to conclude that the chemical does not pose a threat to that very population. In other words, negative epidemiological studies will generally be ignored by the Agency.

- When no adequate human data are available, EPA's default position is that "positive effects in animal cancer studies indicate that the agent under study can have carcinogenic potential in humans." Note that this default is applied in conjunction with the default position concerning negative epidemiological studies addressed in the preceding paragraph. The result is that as little as a single rodent study indicating that a chemical is a rodent carcinogen at high doses will be deemed to trump several negative human epidemiological studies. This is not unbiased, scientific risk assessment seeking to accurately and precisely estimate human health risks. Rather, it is a public policy position masquerading as a risk assessment default

[10] But, of course, there would be no presumption that the substance did pose risks to infants.

procedure. Note that EPA admits as much in the 2003 draft cancer guidelines, referring to the animal study default assumption as a "public-health conservative policy," while acknowledging "the extent to which animal studies may yield false positive indications for humans is a matter of scientific debate" (EPA, 2003a).

- Another EPA default assumption in the case of animal bioassays is that "effects seen at the highest dose tested are appropriate for assessment," although it is "necessary that the experimental conditions be scrutinized." As EPA acknowledges, use of the "maximum tolerated dose" to project effects at low doses raises questions of whether tumorigenic effects seen at high doses are merely the result of cell mortality and regeneration rather than of the substance's inherent carcinogenicity.
- In the case of negative animal bioassays, EPA's default position is as follows: "When cancer effects are not found in well conducted animal cancer studies in two or more appropriate species and other information does not support the carcinogenic potential of the agent, these data provide a basis for concluding that the agent is not likely to possess human carcinogenic potential, in the absence of human data to the contrary" (EPA, 2003a). But, as with negative epidemiological data, EPA hedges on this default, citing its "limitations" and stipulating that because standard bioassays have limited power to detect cancer effects, other information should be considered (e.g., absence of mutagenic or carcinogenic activity among structural analogues) (EPA, 2003a). Again, instead of assessing the weight of the evidence and relying on the result of that assessment, EPA urges its risk assessors to continue to look for any scintilla of evidence that a chemical might possibly have carcinioenic effects and be ready and willing to abandon the weight of the evidence based on that evidence.
- Another EPA default position applied to animal bioassays is that "target organ concordance is not a prerequisite for evaluating the implications of animal study results for humans." This approach is unduly conservative because it calls for extrapolating animal response to humans even in the presence of direct evidence that the response does not occur in humans.

Increase in the IUR for 1,3-Butadiene

In November 2002, EPA updated its IRIS Carcinogenicity Assessment for Lifetime Exposure and Chronic Health Hazard Assessment for Noncarcinogenic Effects for 1,3-butadiene. The support document for this update was EPA's Health Assessment of 1,3-Butadiene (EPA, 2002c). EPA based its cancer assessment on an epidemiology study (Delzell et al., 1996) in which 1,3-butadiene exposure was associated with leukemia in polymer workers exposed to the chemical during styrene-butadiene rubber (SBR) production. Delzell et al. (1996) was a high quality study of workers that showed only a weak association between leukemia and workplace exposures that often were in the range of 10 parts per million – or 1,000,000-fold above the level EPA estimates poses a one in a million cancer risk.

The inhalation unit risk (IUR) represents an estimate of the lifetime extra cancer risk associated with a unit of 1,3-butadiene concentration in ambient air. The new IUR estimate for 1,3-butadiene is 3×10^{-5} $(\mu g/m^3)^{-1}$, which can also be expressed as 0.08/ppm. The corresponding concentration in ambient air that is estimated to pose a one in a million lifetime excess cancer risk is 0.01 part per billion (ppb) (0.03 $\mu g/m^3$).

The new 1,3-butadiene IUR, which is based on a human study, is less stringent than the previous IRIS IUR for this chemical, which was based on an animal study. However, the new IUR is more stringent than EPA's original draft IUR for 1,3-butadiene, which was based on the human study. To understand the changes made by EPA to the IUR for 1,3-butadiene it is helpful to trace the history of the update. This history illustrates two points: (i) how difficult it is for EPA institutionally to make a cancer potency estimate less stringent, even when strong scientific information supports such a decision; and (ii) how EPA's risk assessment methodology can result in an unrealistic estimate of risk if the cumulative impact of the various "health protective" decisions is not carefully evaluated.

Before 1998, the IUR for 1,3-butadiene was 0.7/ppm (2.8×10^{-4} $(\mu g/m^3)^{-1}$), based on tumor data from a mouse bioassay (NTP, 1984). In 1998, EPA issued a draft Health Assessment Document for 1,3-butadiene that proposed a cancer potency estimate of 0.009/ppm (EPA, 1998b). The corresponding level in ambient air estimated to pose a one in a million cancer risk was 0.1 ppb. EPA stated that it had "relatively high" confidence in the excess cancer risk estimate because it was based on "a large, high-quality epidemiologic study [Delzell et al., 1996] in which 1,3-butadiene exposures were estimated for each individual a priori to conducting the exposure-response analysis" (EPA, 1998b). EPA stated further that"[i]t is virtually unprecedented to have such a comprehensive exposure assessment for individual workers in such a large occupational epidemiologic study."

In 1999, EPA made two changes to its cancer potency estimate for 1,3-butadiene in response to recommendations by the EPA Science Advisory Board (SAB) Environmental Health Committee (EHC) (EPA, 1999d). The changes reduced the cancer potency estimate for 1,3-butadiene to 0.0046/ppm, or by about one-half. The corresponding estimate of a level in ambient air that would pose a one in a million excess cancer risk would be 0.2 ppb. This information was

supplied to EPA's Emissions Standards Division for use in certain ongoing regulatory initiatives. EPA advised the Emissions Standards Division as follows:

> The SAB also noted, and we concur, that since the mechanism of action is very different in humans as compared to the mechanism of action in mice, the earlier risk estimate (i.e., IRIS estimate based on mouse tumor data) is not an appropriate basis from which to extrapolate the human risk, hence making the existing IRIS value scientifically unsupportable. The prudent approach for current analysis would be to recognize the Agency's new recalculated 'draft' cancer risk estimate from the current assessment.

(EPA, 1999d). EPA stated further, "This estimate will be presented in our final assessment which is expected to be finished this fall."

Like all EPA cancer risk assessments, the 1999 cancer potency estimate was designed to be an "upper bound" estimate of the potential human cancer risk, with the recognition that the true risk at low exposures could be much lower and could even be zero. There was no scientific reason to believe the 1999 cancer potency estimate was not fully protective of the general population that might be exposed to low levels of 1,3-butadiene in ambient air.

Nevertheless, because the 1999 cancer potency estimate involved lowering EPA's estimate of the potential cancer risk from ambient exposures, it was heavily criticized by some parties. EPA responded to this criticism by extending its deliberations another three years, and finally produced in 2002 a 1,3-Butadiene Health Assessment that combines both human- and animal-based cancer potency estimates (EPA, 2002c). The results of this cancer risk estimate come much closer to the mouse-based risk assessment in IRIS that EPA had said was "not an appropriate basis from which to extrapolate the human risk"(EPA, 1999d).

EPA's deliberations from 1999 to 2002 did not produce a more scientifically defensible cancer risk estimate for butadiene. Rather, EPA's 2002 1,3-butadiene Heath Risk Estimate produced an overly conservative cancer potency estimate for the chemical by departing from normal EPA risk assessment practices, disregarding specific SAB recommendations, and failing to recognize that its numerous "health protective" choices were producing in the aggregate a scientifically implausible result. EPA's primary errors are summarized below:

- <u>EPA ignored SAB advice to adjust for the apparent role of peak exposures</u>. Many of the workers who were studied in Delzell et al. (1996) had very high "peak" exposures to 1,3-butadiene, defined in the study as exposures above 100 ppm. The SAB recommended that EPA exclude those exposures from its dose-response modeling: "In each dose group, adjustment for peak exposures reduced the leukemia risk substantially. Since butadiene exposures to the public will almost never approach the peak exposure range, a more appropriate model for risk would factor out the peak-exposure component" (SAB, 1998). EPA disregarded this advice.
- <u>EPA departed from its usual practice of using the maximum likelihood estimate (MLE) of risk when deriving a cancer risk estimate from human data</u>. In 1998 and 1999, EPA based its

calculations on the MLE of excess cancer risk, which is typical practice for human data, instead of the 95% upper confidence limit (UCL) typically used with animal data. EPA's main reason for use of the MLE in 1998 and 1999 was that "these estimates are based on human data from a large, well-conducted study" (EPA, 1998b, page 9-15). In the 2002 final assessment, EPA switched to the UCL, even though it acknowledged that it "has historically used MLEs for cancer risk estimates from human data rather than upper bounds as used with animal data." (EPA, 2002c, page 10-20) EPA used of the UCL was allegedly based on a "policy position expressed in the 1996 and 1999 proposed [cancer risk assessment] guidelines"(*Id*., page 10-21). However, there is no statement in either document that supports use of the UCL when cancer risk is estimated from human data.

- EPA did not use the model suggested by the SAB for estimating lifetime excess cancer risk in the general population. When estimating the excess cancer risk for the general population, EPA used cumulative lifetime exposure as the relevant measure of dose. The SAB urged instead consideration of a "window of exposure" model that had been used previously by the National Academy of Sciences to estimate lung cancer risk from radon. The SAB stated:

> Regarding the Delzell analysis of butadiene exposure vs. leukemia, … it is noted that 'excluding exposures within 20 years of death weakened and almost eliminated the relationship….' This indicates that in modeling lifetime risk, a model that assumes a limited effect time (i.e., that leukemia risk during a given year of age is affected largely by the butadiene exposures received during the previous, say, 20 years, and only slightly or not at all by more distant ones) should be considered. This 'window of exposure' model has precedents, e.g., lung cancer risk from radon has been modeled in this way in a National Academy of Sciences report ….

(SAB, 1998, page 38) The SAB stated further:

> If this model were considered for projecting lifetime risk, it would show appreciably less risk from chronic exposures than does the present one, which assumes an excess relative risk at, say, age 70 is an additive function of all the exposure accumulated in the previous 69 years.

(*Id*.) There is no scientific basis for believing that exposures to 1,3-butadiene have a significant impact on cancer risk 50, 60, 70 or 85 years later. The data from the Delzell et al. (1996) study and scientific understanding of cancer latency from human studies contradicts such an assumption in this case. EPA's standard approach of using cumulative lifetime dose exaggerates excess cancer risks later in life, and produces an inflated estimate of general population cancer risks. In the case of 1,3-butadiene, a high quality human study supported a different and more scientifically plausible "window of exposure" approach, yet EPA chose not to use that approach, and failed even to respond to the SAB's recommendation.

- EPA's final 2002 Health Assessment computed lifetime excess cancer risks up to age 85, instead of following the Agency's standard practice of calculating risks to age 70. In 1998, EPA calculated lifetime cancer risks up to age 85. The SAB recommended that EPA follow

its normal practice of calculating lifetime cancer risk up to age 70. EPA followed that recommendation in 1999, when it provided an updated cancer potency estimate to the Emissions Standards Division. As EPA explained in its memorandum, "[t]he SAB . . . noted the need to revise the calculations to account for . . . 70 years at risk instead of 85 years" (EPA, 1999d). Nevertheless, EPA's final 2002 assessment offers no explanation for switching back to the calculation based on 85 years. EPA's departure from its typical practice of using 70 years, without explanation, contravenes the core value of "consistency in core assumptions and science policies from case to case" (EPA, 1999b [Cancer Guidelines], page 5-2). Estimating cancer risks to age 85 years also adds to the overstatement of risk caused by EPA's use of lifetime cumulative exposure as the relevant measure of dose.

- EPA departed from its usual practice of basing estimates of lifetime excess cancer risk on general population mortality rates (as opposed to incidence rates). EPA typically derives its estimates of lifetime excess cancer risk for the general population by applying a calculated unit cancer risk estimate (based on human or animal data) to published data on background cancer mortality rates for the general population. EPA does not typically use cancer incidence rates for the general population. In 1998, EPA did not even discuss the possibility of relying on cancer incidence data, and simply took for granted that it would follow the Agency's normal practice of using mortality data. EPA's decision in the final Health Assessment Document to use incidence rates, instead of mortality rates, once again contravenes the core risk assessment value of consistency. No chemical-specific rationale for departing from the standard risk assessment practice was presented.

- EPA applied an extra adjustment factor of 2 to its cancer risk estimate without scientific justification. EPA stated that it applied an extra adjustment factor of 2 to its cancer risk estimate "to reflect evidence from rodent bioassays suggesting that extrapolating the excess risk of leukemia in a male-only occupational cohort may underestimate the total cancer risk from 1,3-butadiene exposure in the general population." (EPA, 2002c, page 10-21) No such factor was applied in the 1998 draft document, nor are we aware of any prior EPA cancer risk assessment where such an adjustment factor has been used.

  EPA asserted that there could be a small excess risk of lung cancer that was not observed in the worker study (EPA, 2002c, pages 10-15 and 10-22). EPA sought to bolster this speculation with a "crude" post-hoc power calculation. However, when making this power calculation, EPA used the MLE of excess lung cancer risk based on female mouse data, whereas elsewhere throughout the document EPA relied on 95% UCL values when deriving cancer risk estimates from animal data. Use of the MLE for the power calculation lowered the estimate of excess lung risk, and thus lowered the power of the study to detect that risk. Use of the MLE just for the "crude" power calculation, and not for any other purpose in the risk assessment document, rendered the entire exercise scientifically suspect.

  EPA's other reasons for adding an additional adjustment factor of 2 were no more persuasive. EPA in effect was combining a human-based cancer risk estimate with a mouse-based estimate, to more closely approximate the latter. EPA stated, "applying a two fold adjustment to the potency estimate of 0.04/ppm derived for leukemia incidence from the occupational epidemiologic study yields a cancer potency estimate of 0.08/ppm, which roughly corresponds to a combination of the human leukemia and mouse mammary gland tumor risk estimates, addressing the concern that the leukemia risk estimated from the

occupational data may underestimate total cancer risk for the general population, in particular females" (EPA, 2002c, page 11-3). Thus, EPA went back to relying on mouse data, despite its earlier recognition that mouse-human differences in mechanism of action were sufficiently great that previous mouse-based risk estimates were "not an appropriate basis from which to extrapolate the human risk" (EPA, 1999d). Moreover, the SAB had concluded that the rat provided a better model than the mouse for human risk assessment. (SAB, 1998, page 36)

- EPA failed to give adequate consideration to the cumulative impact of its many "health protective" choices. The preceding items identify several risk assessment decisions made by EPA that resulted in a substantial overstatement of likely human cancer risks from low level exposures to 1,3-butadiene. EPA's final risk estimate was 20-fold more conservative than the risk estimate it had provided to the Emissions Standards Division in 1999, and only 3-fold less stringent than the mouse-based risk estimate that EPA at that time said was "not an appropriate basis from which to extrapolate the human risk." EPA's Interim Cancer Risk Assessment Guidelines urge "reasonableness" as a core value, and EPA has stated that "common sense and reasonable application of assumptions and policies are essential to avoid unrealistic estimates of risk" (EPA, 1999b, page 5-2). In this case, EPA should have considered whether the collective impact of all of its risk assessment choices was scientifically reasonable. No such analysis was presented in the final health assessment document.
- Finally, because many of EPA's choices were made after the SAB peer review, there was no opportunity for peer review of whether EPA's final risk assessment choices, in the aggregate, were scientifically reasonable. There was no opportunity for external peer review to assess whether EPA's unusual risk assessment decisions (e.g., deciding to switch from the MLE to the 95% UCL, to apply an extra adjustment factor of 2, to ignore the role of peak exposures, to estimate lifetime risks based on 85 years instead of 70, and to use lifetime cumulative dose instead of a "windows of exposure" model as suggested by the SAB) produced a scientifically reasonable result.

The end result of EPA's assessment was an estimate that 10 parts per trillion of 1,3-butadiene in ambient air poses a one in a million cancer risk, even though workplace exposures that often were 1,000,000-fold higher produced evidence of only a weak association between exposure and leukemia, and even though that association almost disappears when "peak" exposures above 100 ppm are excluded. To infer from that data a cancer risk from exposure to 10 part per trillion of 1,3-butadiene in ambient air exceeds the bounds of scientific reasonableness. If there is any cancer risk to the general population from exposure to low levels of 1,3-butadiene in ambient air, it is likely that the upper bound of the estimated risk is well below what EPA presented in its Health Assessment Document.

The 1,3-butadiene cancer risk assessment will play a central role in risk assessments for several categories of stationary sources under the Clean Air Act, and also will be important to ongoing Agency assessments of toxic pollutants from mobile sources. In addition, the general public cannot be expected to understand the conservative nature of EPA's risk assessment. Thus, EPA's excessive conservatism can be expected to have real-world consequences.

Revised RfD for Perchlorate

Ammonium perchlorate is manufactured for use as an oxidizer in solid rocket propellants for rockets, missiles and fireworks. Large volumes of perchlorate have been used in the aerospace and defense industries since the 1950s to fulfill contractual obligations to the Department of Defense (DOD) and NASA, as well as other government agencies. Perchlorate has been found in ground and surface water in 22 states and is pervasive in the Western United States (California, Nevada, Arizona, New Mexico, Texas, Utah and Arizona).

EPA developed an overall model for the perchlorate risk assessment based on perchlorate's mode of action, which is the competitive inhibition of active iodide uptake. In the thyroid gland, iodine is required to produce thyroid hormones. If perchlorate decreases iodine uptake in the thyroid gland, it can eventually lead to decreases in thyroid hormones. The condition of reduced thyroid hormones is called hypothyroidism. EPA has concern that if perchlorate causes maternal hypothyroidism, the developing fetus may be affected. EPA calls the inhibition of iodide uptake the key "event" that precedes hormone and thyroid changes, which in turn could precede neurodevelopmental effects.

In December 1998, EPA published a draft risk assessment recommending a RfD of 0.0009 mg/kg-day (drinking water equivalent of 32 ppb) based on a LOAEL of 0.1 mg/kg-day for thyroid histology and a composite uncertainty factor of 100. (EPA, 1998). An external peer review panel recommended further quality control on histologic endpoints and asserted that the proposed RfD was overly conservative.

In January, 2002, EPA published a draft toxicity assessment for perchlorate, recommending an RfD of 0.00003 mg/kg-day (EPA, 2002). This value equates to a drinking water level of 1 ppb. The draft RfD is substantially more conservative than those previously recommended by EPA (EPA, 1992; EPA, 1995).

There is no supportable scientific basis for the draft perchlorate RfD. To understand the deficiencies in EPA's development of the draft RfD for perchlorate, it is necessary to understand something about the chemistry of perchlorate and its mechanism of action in mammals.

Perchlorate is a negatively charged ion that has the same size and shape as iodide. The perchlorate anion is typically associated with the ammonium cation. Ammonium perchlorate looks, tastes and dissolves in water like table salt.

The thyroid takes in iodide, a necessary nutrient, at the sodium (Na+)-iodide (I-) symporter (NIS) (EPA, 2002). The thyroid uses iodide to make thyroid hormones T3 and T4, molecules that incorporate three or four iodine atoms respectively. The NIS is receptive to perchlorate to a somewhat larger extent than iodide. Thus, when perchlorate is present in the body it interferes with iodine uptake by the thyroid in a dose dependent manner. Inhibition of iodine uptake can be measured by giving a subject radioactive iodine and scanning the thyroid to see how much of the radioactive iodine is absorbed into the thyroid. Inhibition of Radioactive Iodine Uptake (RAIU inhibition) is the only direct effect of perchlorate on mammals.

If iodide uptake is sufficiently blocked for long enough, eventually the production of T3 and T4 will decrease. When it does, the brain releases more Thyroid Stimulating Hormone (TSH) to increase T3 and T4 production. In response to chronically elevated TSH levels, the thyroid gland enlarges (goiter). The most common cause of goiter is iodine deficiency. Iodine is intentionally added to food (mostly table salt) for this reason. There is no evidence of iodine deficiency in the United States.

Presumably, if the TSH concentration is high enough for a long enough time, the thyroid cells are continually stimulated to enlarge and divide. This can lead to mistakes in cell replication and can increase the risk of tumor formation. Blocking of T4 production leading to increased TSH production and thyroid growth is the only cancer mode of action for perchlorate. This has been demonstrated in the rat model but thyroid cancer is not elevated in countries with endemic goiter from iodine deficiency or other dietary goitrogens.

T3 and T4 are necessary to maintain normal metabolism in adults. A pattern of low T4 and high TSH is indicative of hypothyroidism. The most common cause of hypothyroidism world wide is an auto-immune problem where the immune system attacks the thyroid and shuts it down. The treatment for hypothyroidism is taking synthetic T4 in the form of a pill once a day.

T3 and T4 are also needed during fetal development, most importantly for normal development of the brain. In utero, the fetus obtains T3 and T4 from the mother beginning soon after conception. The fetus starts making its own T3 and T4 at the beginning of the second trimester, about the same time as T4 receptors appear in the brain. If a baby is born with a defect such that it cannot produce T3 and T4, it will be normal at birth but will then develop severe mental retardation and certain skeletal defects(cretinism) if he or she is not treated with T4 soon after birth..

In some cases, the thyroid gland is overactive. The most common cause of this is Graves Disease. Starting in the early 50's, perchlorate has been used as a medication to treat overactive thyroid glands. It is still used in Europe for this purpose.

Because perchlorate can reduce T4 production and because reduced T4 concentration can in some circumstances lead to thyroid cancer and defects in brain development in rats, EPA is concerned about the health implications of exposure to environmental levels of perchlorate. However, the available evidence indicates that perchlorate should not be expected to have any adverse health effects on humans at a dose that is at least 200 times higher than the RfD. We discuss the animal and human evidence in turn.

Animal Evidence

As noted above, the draft EPA RfD for perchlorate is based on animal studies in which a LOAEL of 0.01 mg/kg-day was observed (EPA, 2002). The point of departure was based principally on studies performed subsequent to the 1999 external peer review: the "effects study" (Argus Research Labs, 2001); a two-generation rat study (Argus 1999); and a mouse motor activity study performed by the U.S. Navy (Bekkedal 2000). Public comments as well as one member of

the 2002 external peer review panel pointed out that in all of the rodent studies relied upon by the EPA, the principle protein source in the animal diet was soy, a known goitrogen, and that recent studies by NTP and others have demonstrated a profound synergism between soy (isoflavones) and iodine deficiency.

In its "weight of the evidence" assessment, EPA considered several endpoints:

- Motor activity – Based on Bayesian hierarchical analysis of the Bekkedal (2000) (which was reported as a negative study by the authors), combined with a previous study (Argus 1998), EPA determined a LOAEL of 1 mg/kg-day.
- Thyroid tumors – Based on 3 tumors in 2 animals at 19 weeks in first F1 adults (Argus 1999), EPA compared the incidence of all thyroid tumors in NTP archives for rats at 2-year bioassay terminal sacrifice. Applying Bayesian analysis, EPA expressed a "concern" for in utero programming. One member of the 2002 external peer review panel accused EPA of "torturing the data."
- Thyroid histopathology – Based on histopathological findings of hyperplasia, EPA performed a BMDL analysis of the data from the "effects study." The lowest BMDL noted for hyperplasia was 1 mg/kg-day, which was assumed to be equivalent to a NOAEL.
- Thyroid hormones – EPA performed a BMDL analysis of thyroid hormone data (T4) from the "effects" study and determined the lowest BMDL of 0.01 mg/kg-day. Members of the 2002 external peer review panel pointed out that the "statistically significant" hormone changes were well within the normal range and not clinically significant.
- Brain Morphometry – EPA determined a LOAEL of 0.01 mg/kg-day based on statistically significant changes in the size of one brain structure (Argus, 2001). According to EPA, this "point of departure" was selected because when pregnant rats were given this dose during and after pregnancy, pups showed increased widths of some regions in the brain, particularly in the region called the corpus callosum. These effects were seen at only the middle doses given to pregnant rats, not in the control or highest doses. EPA calls this an inverted "U-shaped" dose-response curve. The only neurotoxicologist on the 2002 external peer review panel (Dr. Miki Aschner) reviewed the morphometry data in detail and asserted that the data are un-interpretable and any statistical manipulation of the data therefore meaningless. In comments to the EPA, other neurotoxicologists concluded that the rat brains were sliced in the wrong plane to appropriately evaluate the corpus callosum. EPA nevertheless used the controversial corpus callosum measurements as the "point of departure" for risk assessment (LOAEL 0.01 mg/kg-day).

The LOAEL of 0.01 mg/kg-day was converted to a human equivalent exposure using a physiologically-based pharmacokinetic (PBPK) model. A composite uncertainty factor of 300 was used:

- A three-fold factor for intraspecies variability was used due to the variability observed in the data and PBPK modeling.
- A full factor of ten was applied for LOAEL to NOAEL extrapolation.

- A three-fold factor for study duration was applied due to the concern for the biological importance of the statistically significant increase in thyroid tumors observed in a two-generation reproductive study.
- A three-fold factor was applied, apparently to account for database insufficiency, because "recent studies reinforced concern for [the immunotoxicity] endpoint."

(EPA, 2002d). Thus, the draft RfD is 0.01 x 0.85 / 300 = 0.00003 mg/kg-day.

Human Evidence

In 1952, perchlorate was determined to be more effective than other anions (including nitrate and thiocyanate) in inhibiting iodine uptake by the thyroid. (Stanbury et al., 1952; Wyngaarden et al., 1952, 1953). Subsequently, it has been used as a medication to treat hyperthyroidism associated with Grave's disease. Although more effective treatments for hyperthyroidism have been developed, perchlorate continues to be used medically in some circumstances. Adult dosages of potassium perchlorate of 200 – 900 mg/day produce clinical results.

Employees at Kerr-McGee's Henderson, NV facility were studied (Gibbs et al., 1998) as were employees of Ampac's Cedar City, UT facility (Lamm et al., 1999). Combined results of these studies and BMDL analyses (Crump, 1999) indicate no adverse thyroid or other health effects at dosages up to 0.7 mg/kg-day (DWEL of 25,000 ppb). These employees had worked in perchlorate manufacturing for an average of five years and a maximum of 20 years.

A human volunteer study was done at Boston University with 10mg/day dosing for two weeks and measurement of RAIU inhibition (Lawrence et al., 2000). The researchers noted 40% inhibition of radioactive iodine uptake but no changes in T4 or TSH levels. Although published in the journal *Thyroid*, EPA did not think that the data were useful due to QA/QC concerns.

A second human volunteer study was done in Oregon (Greer et al., 2002) with doses ranging from an equivalent of 200 ppb to 17,000 ppb perchlorate in water. The authors measured RAIU inhibition and thyroid hormones. There was no detectable RAIU inhibition at the low dose (a NOEL) and no hormone effects at the high dose despite 70% inhibition of RAIU. EPA helped design the study for PBPK modeling and used it only for calibration of the rodent data.

Perchlorate occurs naturally in northern Chile. Three coastal cities in northern Chile were located with 110, 6 and ND perchlorate ppb in drinking water. Approximately fifty first grade school children were studied in each city and neonatal screening data for a three year period from the same three cities were evaluated (Crump et al., 2000). There were no adverse thyroid or any other health differences attributable to life long exposure to perchlorate at 110 ppb. Serum and urine perchlorate levels among the school children drinking water with 110 ppb were consistent with the water level (27.5 kg child drinking 1 liter per day).

Thyroid hormone concentrations were compared for infants born to consumers of Las Vegas, Nevada, drinking water (which contains approximately12 ppb perchlorate) and infants born to consumers of Reno, Nevada, drinking water (no perchlorate detected in three published studies). No differences in neonatal thyroid screening T4 or TSH results (Li et al., 2000, Xiao et al., 2000)

or Medicaid data regarding prevalence of thyroid diseases or thyroid cancer were found (Li et al., 2001). There is no increase in neonatal hypothyroidism in southern California in zip codes associated with elevated perchlorate exposure (Lamm & Doemland, 1999).

An unpublished Masters thesis from Berkeley found a dose-related difference in neonatal thyroid hormones. (Schwartz, 2001). Faculty at the School of Public Health at Berkeley have recently performed a similar study and obtained negative results (Kelsh et al., 2003).

A study comparing neonatal screening data from Yuma, Arizona (6 ppb perchlorate in drinking water) and Flagstaff, Arizona (no detectable perchlorate in drinking water) found a slight difference in TSH (Brechner et al., 2000; Crump et al., 2001; Goodman 2001). The Yuma population has been revisited and the approximate half of the population with perchlorate exposure were found to have similar thyroid hormone levels as the approximate half of the population with no perchlorate exposure. Most of the difference in neonatal TSH levels between the Yuma and Flagstaff poputations was attributable to infant age at the time of testing.

EPA critiqued the human data in the 2002 draft risk assessment (EPA, 2002d). Other than a discussion of particle size of dust in the occupational studies (irrelevant because serum and urine levels confirmed absorption), nearly all of the critique of the human studies was by a single author (Park, 2001) from NIOSH who was apparently contracted by EPA.

* * *

The human studies provide a logical framework for deriving an RfD that is inherently protective and that renders uncertainty factors superfluous. Together, these studies show that for any adverse effect (e.g., clinical hypothyroidism) to occur, the dose of perchlorate would need to be high enough to cause significant (probably greater than 70%) inhibition of iodine uptake for several years (Lamm et al, 1999).

Based on Greer et al. (2002), the RfD should be 0.005 to 0.17 milligrams per kilogram of body weight per day, equivalent to 175 to 6,000 parts per billion in drinking water. The bases for the uncertainty factors that should be applied to the Greer study are as follows:

- LOAEL to NOAEL uncertainty factor: a value of less than 1 (e.g., 0.1 to 0.01) is used since the study provides an estimate of a no observed effect level (NOEL) instead of a no observed *adverse* effect level (NOAEL).
- Interspecies uncertainty factor: no extrapolation from animals to humans due to use of human data; therefore, 1 is appropriate.
- Intraspecies uncertainty factor: 3 to 10 for use of healthy adults in the study.
- Database uncertainty factor: the perchlorate database is extensive, including several studies in human populations, and its effects have been well characterized; therefore, 1 is appropriate.
- Subchronic to chronic uncertainty factor: due to the implausibility of chronic effects from perchlorate in the absence of acute effects, a factor of 1 is appropriate.

Proposed RfD for Acetone

EPA posted a draft Toxicological Review and draft IRIS summary for acetone on its web site on August 16, 2001, at the same time the documents were provided to external peer reviewers. EPA has proposed an oral RfD for acetone of 0.3 mg/kg/day. This value is more than 100-fold below normal endogenous production of acetone in healthy individuals. EPA reached this result by applying "standard" uncertainty factors that are not scientifically appropriate. EPA also understated the amount of information available to evaluate potential hazards from exposure to acetone, resulting in application of an additional uncertainty factor that further skews its RfD. EPA toxicity estimate for acetone is notably inconsistent with those of other scientists.

Acetone is naturally present throughout the human body as a result of its production during fatty acid catabolism. Infants and young children typically have higher acetone blood levels than adults due to their higher energy expenditures. Vigorous exercise, dieting, pregnancy, and lactation can also lead to normal fluctuations in the blood levels of acetone without any ill effect. The rate of acetone production in normal healthy adults is approximately 41 mg/kg/day (equivalent to approximately 2.9 g/day).[11] Thus, the proposed oral reference dose (RfD) for acetone of 0.3 mg/kg/day is more than 100-fold below normal endogenous production of acetone in healthy individuals. A daily dosage in the magnitude of the RfD is meaningless from a toxicological perspective, given endogenous production levels.

EPA's proposed RfD for acetone is also inconsistent with the toxicity assessments performed by other scientists and groups. The external co-author of the draft IRIS Toxicological Review, Dr. Forsyth of Oak Ridge National Laboratory, recommended an RfD for acetone of 3.0 mg/kg/day. In addition, the World Health Organization (WHO) has published an Environmental Health Criteria document for acetone that contains a recommended value of 9.0 mg/kg/day – a value that is 30-fold above EPA's recommendation (WHO, 1998). The WHO value is still below normal endogenous production rates in healthy individuals, but it is more scientifically plausible than the value proposed by EPA. The values differ from EPA's because both Dr. Forsyth and WHO use more scientifically defensible uncertainty factors than were applied by EPA.

Acetone exhibited very low toxicity in 90-day drinking water studies sponsored by the National Toxicology Program (NTP). Minimally toxic concentrations were estimated to be 20,000 ppm (1,700 mg/kg/day) for male rats, 20,000 ppm (4,858 mg/kg/day) for male mice, and 50,000 ppm (11,298 mg/kg/day) for female mice. No toxic effects were identified in female rats at the highest concentration of 50,000 ppm (3,100 mg/kg/day). NTP recommended against the conduct of chronic studies of acetone because "the prechronic studies only demonstrated a very mild toxic response at very high doses in rodents," and because of "the absence of any evidence supporting the carcinogenic potential for acetone" (NTP, 1989).[12] In other words, no chronic

[11] See G. A. Reichard et al., Plasma acetone metabolism in the fasting human. J. Clin. Invest. 63, 619-626 (1979), cited in Table 74.33 in Patty's Toxicology, Fifth Edition, Volume 6, Edited by Eula Bingham, Barbara Cohrssen, and Charles H. Powell.

[12] This recommendation was adopted by the Hazardous Waste Information Evaluation Subcommittee (HWIES) of

toxicity/oncogenicity study has been conducted for acetone because acetone exhibits such low toxicity that NTP has concluded chronic toxicity studies are not necessary.

In deriving the proposed RfD, EPA applied a combined total uncertainty factor of 3000 -- factors of 10 were applied for intraspecies extrapolation, subchronic-to-chronic extrapolation, and "database insufficiency", including the absence of a chronic study, and a factor of 3 was applied for interspecies variability. These factors are clearly excessive.

EPA did not provide a justification for applying an uncertainty factor for subchronic-to-chronic extrapolation, apart from noting that this is a "standard factor," and speculating that repeated exposures over an extended period of time "could lead to more pronounced effects." The available scientific data, however, do not support application of a factor of 10, and strongly contradict EPA's speculation. As noted above, NTP expressly concluded that the effects of acetone were so mild, at such high doses, that a chronic study was not necessary. Further, several published studies support the use of an uncertainty factor of less than 10 for the absence of a chronic study, particularly for substances like acetone that are readily metabolized and eliminated from the body. See, e.g., Dourson et al. (1996), Beck et al., (1992) and Nessel et al. (1995). Thus, the subchronic-to-chronic UF applied by EPA clearly is overly conservative in light of the available data on acetone.

The UF of 10 for database insufficiency also is excessive. EPA takes an overly compartmentalized approach, and fails to make use of inhalation studies of acetone (which also demonstrate low toxicity) when deriving the oral reference dose. In addition, isopropanol has been shown to be extensively metabolized to acetone and several TSCA guideline studies for this compound are available. The combination of inhalation data for acetone and numerous studies on isopropanol (including studies of chronic toxicity, neurotoxicity, developmental toxicity, reproductive toxicity and developmental neurotoxicity) are ample to fully evaluate the potential toxic effects of acetone, rendering a UF for database insufficiency completely unnecessary. EPA's draft IRIS summary drastically understates the amount of scientific information available to evaluate potential hazards from exposure to acetone. The draft IRIS summary gives a "low" rating to the database, but when acetone went through the OECD "Screening Information Data Set" (SIDS) review process, with the United States as the sponsor country, the SIDS Initial Assessment Report (SIAR) concluded that "[t]he human health and environmental effects of acetone have both been well studied" (EPA, 1999h). The SIAR reported that the most significant health effects of acetone are eye irritation and "an acute effect on the central nervous system," but noted that "high exposures are required and health hazards are slight," making acetone "a low priority for further work" (EPA, 1999h). Thus, the Council believes that no factor for database insufficiency is justified.[13]

---

the Public Health Service Committee to Coordinate Environmental Health and Related Programs. The recommendation of HWIES in turn was accepted by the Agency for Toxic Substances and Disease Registry (ATSDR), which had been considering proposing acetone for possible chronic toxicity testing. See 54 Fed. Reg. 42042 (October 13, 1989); 55 Fed. Reg. 34966 (August 27, 1990).

[13] The SIAR found that acetone has "low potential for systemic toxicity" and "showed minimal reproductive and developmental effects in animals exposed either by inhalation or via drinking water." The SIAR concluded that "acetone does not pose a neurotoxic, carcinogenic, or reproductive health hazard at the concentrations found anywhere in the environment." Indeed, the SIAR posits that the "ability of humans to naturally produce and dispose

An overly conservative approach to deriving an oral RfD can have significant negative consequences. By applying excessive uncertainty factors to materials of demonstrated low toxicity such as acetone, EPA makes it harder for prospective users of IRIS to make rational distinctions among compounds, and therefore harder to manage potential hazards and risks as effectively as possible. Further, when excessively conservative IRIS values are applied in particular regulatory settings, potential hazards may be identified where in fact none may reasonably be anticipated, and substantial resources may be wasted addressing scientifically implausible risks.

Moreover, the issues associated with the acetone RfD are compounded because, although the IRIS process has included an external peer review, EPA has refused to provide the external peer reviewers with the comments submitted by interested parties. IRIS documents are intended to serve as the starting point for risk assessments conducted by EPA program offices. EPA also expects its documents to be used by other federal and state agencies, and by other stakeholders and the general public. Given the breadth and importance of EPA's IRIS documents, peer reviewers should have access to all relevant information, including scientific input provided by interested parties (especially if it differs from what is contained in the draft documents). To command respect in regulatory and scientific communities, EPA's IRIS files must be the product of an open and unbiased process, where public comment is encouraged and all comments are given fair consideration. By not providing public comment to peer reviewers, the Agency undermines the validity of its findings.

---

of acetone may to a large degree explain its relatively low toxicity following external exposure to moderate amounts of the vapor or liquid." The SIAR was approved in its entirety by EPA scientists. Compared to the draft IRIS summary, the SIAR provides a more realistic and balanced assessment of the adequacy of existing data and acetone's potential health hazards.

Proposed RfD for Trichloroethylene

In developing an RfD for trichloroethylene (TCE), EPA applied, without adequate justification, several extremely conservative uncertainty factors. The result is an RfD that is one to two orders more stringent than necessary.

EPA identified several toxic effects associated with TCE exposure including liver, kidney, and developmental effects. It appears that the lowest doses are linked to changes in the liver weight to body weight ratio in both mice and rats. EPA (2001d) reports these doses in "human-equivalent terms" based on pharmacokinetic modeling performed by Clewell et al. (2000) and Barton and Clewell (2000). EPA selected a human-equivalent dose of 1 mg/kg-day as the point of departure, supported by liver toxicity observed in three studies (Tucker et al., 1982; Buben and O'Flaherty, 1985; Berman et al., 1995).

EPA characterized uncertainty associated with the reference values by applying several uncertainty factors to the "point-of-departure" dose. For the RfD, EPA (2001d) assigned a value of 50 for human variation and values of 3 for animal to human extrapolation, subchronic to chronic exposure, LOAEL to NOAEL extrapolation and background exposures, resulting in an overall uncertainty factor of 5000. This overall factor exceeds EPA's own maximum composite factor of 3000. Thus, EPA (2001d) limited the uncertainty factor to 3000, and applied it to the point of departure dose of 1 mg/kg-day to derive an oral RfD of 3×10^{-4} mg/kg-day for TCE.

Human Variation

EPA (2001d) suggests the application of a 50-fold safety factor to account for human variation, based on a 3-fold factor to account for human pharmacodynamic differences and a 15 to 20-fold factor to account for uncertainties in the pharmacokinetic models applied to estimate human doses (this latter factor also subsumes the pharmacokinetic uncertainty associated with extrapolation from animals to man). EPA (2001d) justifies the 15 to 20-fold factor as the span between the 50th and 99th percentile of a range of potential values for two dose metrics arbitrarily modeled as log-normal.

Because 50% of the possible values for the dose metric are below the median, the 15 to 20-fold factor beyond the median may be characterized as a value that is applicable to, or too high for, 99% of the possible cases. Because of the (arbitrary) log-normal nature of the distribution, the "multiplier" on the median dose metric falls very quickly. Thus, the 95th percentile of the dose metric is more than 50% lower (i.e., if one selected the 95th percentile for determining the magnitude of the uncertainty factor, the value would be 7 to 8, rather than 15 to 20). The 90th percentile is lower still (i.e., the factor at this percentile would be approximately 4). As such, it would appear that this unprecedented high uncertainty factor proposed by EPA would provide very little additional protection;.

One must also use caution with a modeled distribution in that the extremes may be a statistical aberration rather than a duly conservative value from the upper range of a population of actual empirical values. By analogy, in dealing with uncertainty in the exposure assessment component

of the risk paradigm, EPA generally suggests a value with 90 to 98 percent confidence (EPA, 1992a). This guidance represents overall confidence, not the confidence derived from a single selected factor. Thus, attempting to achieve 99% confidence from a single uncertainty factor where other factors are also being used is excessive and inconsistent with typical regulatory practice. As such, it is suggested that an uncertainty factor to account for human pharmacokinetic variation be no greater than 4 (representing the 90th percentile on the dose metric as a multiple of the median) rather than 15 to 20. Multiplying this 4-fold human pharmocokinetic variation by an uncertainty factor of 3 for human pharmacodynamic variation results in an overall uncertainty factor for human variation of 10 when rounded to one significant digit. .

Furthermore, in their discussion of the intraspecies uncertainty factor, the TCE Reference Dose Technical Panel (EPA, 2001e) reported that Renwick and Lazarus (1998) demonstrated that an uncertainty factor of 10 accounted for "variability in both kinetics and dynamics in the vast majority of the population (>99%)."

Subchronic to Chronic Dosing

EPA (2001d) applied an uncertainty factor of 3 to account for subchronic to chronic exposure. EPA mischaracterized the studies providing a point of departure for liver endpoints as "subchronic." Tucker et al. (1982) exposed animals for 6 months. This is well in excess of rodent dosing periods typically characterized as "subchronic."

The definition of subchronic in the IRIS glossary is 10% of an animal's lifetime, which would be approximately 90 days. The EPA Office of Prevention, Pesticides, and Toxic Substances, OPPTS Protocol 870-3100, also characterizes a 90-day testing protocol as a "subchronic" test. And, the Reference Dose Technical Panel (EPA, 2001e) proposes the following definition for chronic exposure when establishing a chronic reference value: "repeated exposure by the oral, dermal or inhalation route for more than approximately 10% of the life span in humans." Furthermore, standard toxicology references typically characterize chronic dosing as greater than 10% of the test species' lifetime (e.g., Stevens and Mylecriane, 1994).

Moreover, Barton and Clewell (2000) states clearly that no uncertainty factor is needed to account for subchronic to chronic exposure. The authors report:

> Selection of changes in liver weight/body weight as a potential critical endpoint was based on its role as an early event in the toxicity process and a sensitive indicator of potential liver effects observed at later times. Therefore, based upon the mode-of-action argument that this early event is an indicator of toxicities that develop later, no adjustments for the duration of exposure would be needed, regardless of the study duration.

Finally, in 1998, the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use indicated that little difference was seen in toxicity in rodents dosed for 6 months rather than one year and advocated that 6-month dosing studies in rodents be specified as chronic (ICH, 1998). European Union countries have adopted

this policy. Thus, the Tucker et al. (1982) study should be considered chronic and not subject to an uncertainty factor for adjustment from subchronic to chronic dosing.

LOAEL to NOAEL

EPA (2001d) applied an uncertainty factor of 3 to correct for potential differences in dose required to adjust the LOAEL to a NOAEL. However, in one of the three liver studies used (Tucker et al., 1982), the endpoint was in fact a NOAEL. In another (Berman et al., 1995), an LED10 was calculated. This metric is frequently used in a fashion similar to a NOAEL i.e., an uncertainty factor for LOAEL to NOAEL is infrequently applied to the LED10 (Faustman, 1996). In fact, the Reference Dose Technical Panel (EPA, 2001e) asserts that the LOAEL-NOAEL uncertainty factor is "unnecessary when using dose-response modeling to derive a benchmark dose, as the value at a given level of response can be derived from the dose-response model."

Finally, the endpoint of the critical studies subject to this uncertainty factor was change in the liver weight to bodyweight ratio. In all cases, the change in liver weight was not accompanied by histopathological or chemical indications of injury. It is disturbing that EPA treats these data as if they were clearly indicative of an adverse effect. Barton and Clewell (2000) clearly describe their interpretation of the nature of the liver effects as a justification for applying no LOAEL to NOAEL uncertainty factor.

Animal to Human

EPA (2001d) applied an uncertainty factor of 3 to account for animal to human extrapolation. This factor covers animal to human pharmacodynamic variation (i.e., species sensitivity).

Barton and Clewell (2000) state that "information on the mode of action does not support the default assumption that humans are more sensitive than animals for liver effects." Liver effects associated with TCE exposure likely involve the peroxisome proliferator-activated receptor (PPAR). Available data indicate that mice have a fully active PPAR, whereas rats and humans are less responsive (Barton and Clewell, 2000). This suggests that humans are not more sensitive than the most sensitive rodent. Indeed, Barton and Clewell (2000) report "these data indicate that the value of the uncertainty factor for interspecies extrapolation should be no greater than 1 and potentially less." The Reference Dose Technical Panel (EPA, 2001e) reports that an uncertainty factor of 3 accounts for pharmacodynamic differences, but this default factor should be adjusted when "data support the conclusion that the test species is more or equally as sensitive to the pollutant as humans."

"Modifying Factor" for Background Exposures

EPA (2001d) used a so-called "modifying factor" of 3 to reflect background exposures to TCE and TCE's metabolites and thereby protected against cumulative risk. EPA's use of a modifying factor for this purpose is unprecedented. The issue of cumulative risk is a risk management rather than a risk assessment issue (NRC, 1983).[14]

[14] Note also that the need to assess background sources of TCE is unclear. Wu and Schaum (2000) indicated that

Finally, note that the use of modifying factors for any purpose is disfavored. The Reference Dose Technical Panel (EPA, 2001e) considers "the modifying factor (MF) to be sufficiently subsumed in the general database uncertainty factors." Furthermore, the Panel considers "the availability of a factor that may be evoked with quantitative consequences based solely on professional judgment or assessment as being counter to the other stated intentions of risk characterization, i.e., that decisions and procedures within assessments reflect clarity, transparency, consistency and reasonableness." The Panel therefore recommended discontinuing the use of modifying factors.

* * *

Based on review of the various uncertainty factors, EPA's uncertainty factor for human variation should be reduced by at least a factor of 2 and the modifying factor should be eliminated. There also appears to be little justification for EPA's uncertainty factors for animal-to-human extrapolation, LOAEL-to-NOAEL extrapolation, and subchronic-to-chronic adjustment. Thus, EPA's overall uncertainty factor is at least an order of magnitude and more likely two orders of magnitude greater than scientifically reasonable. This recommendation is strongly supported by the RfDs derived in Barton and Clewell (2000).

ambient concentrations of trichloroethylene decreased significantly between 1987 and 1994. As such, the exposure estimates provided in EPA (2001d) are likely to be incorrect for the current situation.

NTP Proposed Listing of Naphthalene as a Carcinogen

The National Toxicology Program (NTP) has proposed to list naphthalene in the Eleventh Edition of the Report on Carcinogens as "reasonably anticipated to cause cancer in humans." As discussed below, the proposed listing is based on, but ignores, NTP's own criteria for listing chemicals as possible carcinogens.

In order to be listed as "reasonably anticipated to cause cancer in humans" by NTP, a compound generally must first meet one of two criteria. Either there must be:

> limited evidence of carcinogenicity from studies in humans which indicates that causal interpretation is credible, but that alternative explanations, such as chance, bias, or confounding factors, could not be adequately excluded,
>
> or
>
> sufficient evidence of carcinogenicity from studies in experimental animals which indicates there is an increased incidence of malignant and/or a combination of malignant and benign tumors.

If the listing is based on "sufficient evidence" in animals, another criteria must also be met: the animal data must be "in multiple species or at multiple tissue sites; by multiple routes of exposure; or to an unusual degree with regard to incidence, site, or type of tumor or age at onset." Finally, if a chemical does not have sufficient evidence of carcinogenicity in animals, NTP can list the chemical as "reasonably anticipated to cause cancer in humans" based on other considerations of structure and mechanism.

NTP listing criteria require that all conclusions be made based on scientific judgment with consideration of dose-response, metabolism, pharmacokinetics, and other relevant information. NTP specifically states: "substances for which there is evidence of carcinogenicity in laboratory animals are not considered 'reasonably anticipated to cause cancer in humans' where there are compelling data indicating that the agent acts through mechanisms which do not operate in humans."

Based on the NTP's own criteria, NTP should not list naphthalene as "reasonably anticipated to cause cancer in humans" for two reasons. First, there is no "limited evidence of carcinogenicity from studies in humans" reported in the literature. The only human studies that have been discussed by NTP are very small health status surveys of employees of an East German coal tar company "engaged in the purification of naphthalene" conducted in the early 1970s (Wolf, 1976, 1978). Although these reports have some relevance to "naphthalene workers" because naphthalene is a component of coal tar, Wolf (1978) suggested that tar fumes in combination with heat were causative factors in the development of laryngeal cancer. In addition, the study was seriously confounded in that four of the 15 workers in the study developed laryngeal cancer and all four were smokers. In addition, all of the workers were likely to have had many confounding chemical exposures, several of which were discussed by Wolf (1976). As noted in

comments prepared by AMEC[15] on the NTP proposal to list naphthalene as a carcinogen (Appendix 5), the vast literature on the health status of thousands of workers in numerous industries who were exposed to naphthalene-containing mixtures reveals no indication that naphthalene exposure was responsible for an increase in cancer rate. Moreover, even in studies involving workplace exposure to multiple chemicals, nasal tumors, the only tumor type associated with naphthalene exposure in rodents, were not elevated. Clearly, there is not "limited evidence of carcinogenicity from studies in humans" for naphthalene.

Second, under the NTP criteria there is not "sufficient evidence of carcinogenicity from studies in experimental animals" for naphthalene. The evidence of carcinogenicity is only in one species, not multiple species; the evidence is at one tissue site, not multiple sites; the evidence is from one route, not multiple routes of exposure; and the evidence does not show an unusual high incidence or suggest that site, tumor type or age at onset were in any way unusual. Moreover, there are "compelling data indicating that the [naphthalene] acts through mechanisms which do not operate in humans." Specifically, human nasal physiology is differs significantly from rodent nasal physiology. A primary site of action for toxic effects in rats is the olfactory epithelium, which comprises a significant portion of the total nasal cavity. The rat is an obligatory nose breather and must rely on olfaction for survival. The olfactory mucosa in rats is a highly developed system of cellular structures that performs complicated integration of olfaction and air humidification. Approximately 50% of the total surface area of the posterior region of the rat nasal cavity is composed of the olfactory epithelium (Gross et al., 1982; Uraih and Maronpot, 1990). Inhaled vapors need traverse only a few millimeters past the resistant respiratory epithelium to reach the sensitive olfactory tissue in rats.

By comparison, the total surface area for chemical exposure is much less in humans (by a factor of five) since human nasal turbinates are much less convoluted than in the rodent. The olfactory epithelium comprises only about 10% of the human nasal cavity and is confined to the posterior dorsal region of the nasal cavity (Frederick et al., 1994). The ciliated respiratory epithelium is the major lining of the human nasal cavity. In humans, inhaled vapors must traverse several centimeters through the ciliated respiratory epithelium before reaching the olfactory epithelium. Through mucociliary actions, the respiratory epithelium provides a protective system for the olfactory epithelium and other respiratory tissues. As a result of these differences, the efficiency of extracting chemicals from air inhaled through the nose is much less in humans than in rodents, which rely heavily on their sense of smell to locate food. The resulting dose deposited to the human olfactory epithelium, in particular, from inspired air is far less than for rodents for any given naphthalene concentration in air.[16]

It is therefore clear that the mechanism of action in the rat is not relevant to the human, and this fact alone should require NTP to conclude that naphthalene does not meet the criteria for listing. Moreover, the empirical evidence showing a lack of nasal tumorigenic response in humans is

[15] AMEC. 2003. Comments on Proposal to List Naphthalene in the Report on Carcinogens, Eleventh Edition. Submitted to the National Toxicology Program. March 21.

[16] While the paragraphs above focus on physiological differences relevant to understanding the differences between rat and human upper respiratory tract mechanisms of actions, it is also important to note that recent NIH-funded research demonstrates that the metabolism of naphthalene differs between rodents and humans. See, e.g., Buckpitt et al.. (2002).

entirely consistent with the lack of relevance of the mechanism of action of high dose naphthalene exposure in rats to the human situation. Thus, there are "compelling data indicating that the agent acts through mechanisms which do not operate in humans."

EPA's New RfC for Naphthalene

EPA has adopted an IRIS Reference Concentration (RfC) for naphthalene 0.003 mg/m^3. This value is slightly lower than the ambient background concentration for naphthalene (0.0052 mg/m^3) in the United States (ATSDR, 1995). EPA's RfC for naphthalene therefore suggests that a substantial portion of the United States population faces health risks from exposure to naphthalene. In fact, as discussed below, there is almost certain no risk whatsoever from exposure to ambient levels of naphthalene. Not even workers exposed to occupational levels thousand of times higher than EPA's RfC are at significant risk.

EPA's RfC for naphthalene is based on nasal irritation in mice which was seen at a LOAEL of 50 mg/m^3. EPA derived its RfC by applying an uncertainty factor of 3000 to a human equivalent LOAEL of 9.3 mg/m^3. The total uncertainty factor of 3000 is comprised of factors of 10 to extrapolate from mice to humans, 10 to protect sensitive humans, 10 to extrapolate from a LOAEL to a NOAEL, and 3 for database deficiencies (IRIS, 2003).

As discussed in the previous section of these comments, using nasal effects in rodents to assess the risks of human exposure to naphthalene is inappropriate because humans are less sensitive to inhaled naphthalene than mice. Even more important, there is no need to rely on animal data to develop a naphthalene RfC for humans because several worker studies exist that can be used to determine a safe inhalation exposure level for humans (ACGIH, 1993; OSHA, 1995).

In fact, OSHA has relied on these studies to determine a time-weighted average threshold limit value (TWA-TLV) or permissible exposure level (PEL) of 50 mg/m^3 (10 ppm). This limit protects workers from significant risks of eye irritation and other ocular effects from exposure to naphthalene. Note that this value is five times higher than EPA's starting point for deriving the RfC (the human equivalent LOAEL of 9.3 mg/m^3 from the mouse study) and over 16,000 times higher than EPA's RfC.

Hoboken, New Jersey, Industrial Building Remediation

When it came to EPA's attention that an industrial building in Hoboken, New Jersey, had been illegally converted to residential use, the building was vacated and EPA sought to remediate the building under Superfund. Although it was generally agreed that the building should not be used for residential purposes after remediation, the owner proposed to continue to use it as a factory after remediation. EPA ultimately adopted a mercury remedial objective that was so stringent that it effectively required demolition of the building rather than re-use for industrial purposes.

The EPA remedial objective -- 0.44 µg of mercury per cubic meter of air – was unduly stringent because:

- EPA ignored the fact that no other organizations charged with worker protection have adopted an airborne mercury standard that is anywhere as stringent as EPA's objective. Table 2 demonstrates this.

Table 2: National and International Occupational Standards for Elemental Mercury

Standards Organization	*Mercury Air Standard[17] (µg/m^3)*
ACGIH (US)	25 (skin)[18]
NIOSH (US)	50 (skin)
OSHA (US)	100 (skin)[19]
Australia	50 (skin)
Belgium	100 (skin)
Canada	50
China	20
Egypt	50
Finland	50
France	50 (skin)
Germany	100 (skin)
Hungary	20[20]
India	50

[17] 8-hour Time-Weighted Average ("TWA").

[18] "Skin" notation indicates that standard is intended to protect against both inhalation and dermal absorption of mercury.

[19] A revised standard of 50 µg/m^3 was struck down in AFL-CIO v. OSHA, 965 F.2d 962 (11th Cir. 1992) on procedural grounds.

[20] "Target" value.

Mexico	50
Poland	50
Sweden	50 (skin)
Switzerland	50 (skin)
Turkey	100 (skin)
United Kingdom	25
World Health Organization	25

- EPA's standard was derived by starting with the air concentration generally believed to be protective (25 µg/m^3) and, in effect, applying a total safety factor of over 50 to reach its objective of 0.44 µg/m^3. Although stated to be intended to protect workers from airborne exposure to mercury, the objective was in fact derived using residential exposure assumptions (exposure 24 hours a day, seven days per week).
- EPA simply ignored the evidence that the 25 µg/m^3 standard is protective for workplace exposure. Specifically:
    - The most reliable worker exposure studies (Fawer et al.., 1983; Piikivi et al.., 1989a,b,c) concluded that an occupational limit of 25 µg/m^3 is adequately protective.
    - Twenty-one studies have found that adverse effects in humans occur only at mercury exposure levels resulting in urinary concentrations above 50 µg mercury per gram of creatinine. This concentration per gram of creatinine corresponds to 110 µg of mercury per liter of urine, which in turn corresponds to an airborne exposure level of about 41 µg/m^3.

PCB TSCA "Megarule"

In June 1998, despite its own conclusion that the appropriate CSF for PCBs was in the range of 0.1 to 2.0 $(mg/kg/day)^{-1}$ (EPA, 1996a), EPA promulgated the so-called "PCB Megarule" based on a 4.0 $(mg/kg/day)^{-1}$ CSF. EPA doubled the 2.0 CSF to take into account what it conceded were unquantified "non-cancer" effects of PCBs:

> The cancer slope factor used in the risk estimates generated in this report is 4.0 $(mg/kg/day)^{-1}$. This is the value used by USEPA's Office of Toxic Substances to evaluate risks for USEPA's PCB Spill Cleanup Policy and was also requested for use here by USEPA's Office of Pollution Prevention and Toxics (OPPT) (personal communication between John Smith, USEPA and Linda Phillips, Versar, Inc., August 21, 1997) as a conservative value that would account for uncertainty associated with the presence of potentially toxic dioxin-like congeners and other uncertainties associated with the exposure/risk assumptions. This value is somewhat more conservative than the values listed in recent Agency guidance (USEPA, 1996a) and used in some of the risk assessments generated by other agencies.

USEPA, Assessment of Risks Associated with the PCB Disposal Amendments (Versar, May 11, 1998). In response to comments asking EPA to base the Megarule on the latest scientific information, EPA stated in the preamble to the final rule:

> In adopting this policy position, EPA weighed the potential benefits and costs associated with revising the final rule to reflect the most recent PCB cancer potency information. Such a change at this time would delay the issuance of the final rule and its anticipated large cost savings, for likely only very marginal benefits.

63 Fed. Reg. 35383, 35386 (June 29, 1998). EPA's "Response to Comments" background document stated:

> While the 4.0 $(mg/kg/day)^{-1}$ slope factor does not correspond with any of the cancer slope factors in the September 1996 report [the Reassessment], it does allow for additional protection from as yet unquantified risks from non-cancer human health effects and effects to the environment.

USEPA, Response to Comments Document on the Proposed Rule -- Disposal of Polychlorinated Biphenyls (May 1998) at 132.

Industry challenged the Megarule, arguing that the 4.0 $(mg/kg/day)^{-1}$ CSF could not be defended based on the record or science. Industry pointed out that it is well-known -- and accepted by

EPA and all other entities to assess risk from chemical exposure -- that cancer and non-cancer risks are not summed in the course of risk assessment. Rather, cancer and non-cancer risks are estimated separately and the more stringent of the risk estimates controls.

After lengthy settlement negotiations, EPA agreed to a remand of the 4.0 $(mg/kg/day)^{-1}$ CSF. The Court then remanded the matter to EPA. Central & Southwest Services, Inc. v. EPA, 220 F.3d 683 (5th Cir. 2000). To date, EPA has not proposed corrected Megarule standards for PCBs.

Gas Turbine Association Petition to "Delist" Gas Turbies from MACT

As discussed above, the Gas Turbine Association has petitioned EPA to "delist" gas turbines from MACT requirements pursuant to Section 112(c)(9) of the Clean Air Act on the grounds that gas turbine emissions present less than a 10^{-6} risk of cancer and an insignificant non-cancer risk. As previously related, USEPA is demanding that GTA use very conservative exposure assessment parameters in the risk assessment supporting the delisting petition. In addition, EPA is requiring use of a scientifically unsupportable approach to toxicity assessment

EPA is insisting that GTA's risk assessment assume that the risk from all of the carcinogenic HAPs is additive. This is inconsistent with the accepted risk assessment practice that assumes risk to be additive only when the chemicals at issue affect the same target organ or have the same mechanism of action. EPA (1986) cautions that "if the compounds in a mixture do not have the same mode of toxicologic action, dose additivity is not the most biologically plausible approach, and can lead to substantial errors in risk estimates if synergistic or antagonistic interactions occur (emphasis added)." Other agencies express similar concerns. The Scientific Committee on Food (SCF, 2002) in its research on food toxins cautioned that dose additivity requires a common mechanism of action. In addition, the Department of Energy (DOE, 1995) in its Reference Manual for CERCLA's Baseline Risk Assessment advises that "if two or more components each act by different toxicological mechanisms, additivity of risks for a common endpoint is not necessarily to be expected."

EPA understands well that it would not be appropriate to assume additivity without substantial evidence. The Food Quality Protection Act (FQPA) requires the assessment of cumulative risks that might result from exposure to pesticides and other substances that are toxic by a common mechanism. In compliance with the FQPA, EPA has set forth specific guidelines for identifying those substances that have a common mechanism of toxicity (EPA, 1999e). Those guidelines include a thorough identification and analysis of all information that can be used as the basis for determining the mechanisms of toxicity for each pesticide and a weight-of-evidence approach to support the common mechanisms of toxicity. EPA (1999e) states that "no single piece of information will suffice to support the characterization of a specific or common mechanism of toxicity; this finding will be supported by the analysis and inter-relationships of available pieces of information." In developing its guidelines, EPA (1999e) has been careful not to confuse "mechanisms of toxicity" with "site of toxic action" or "site of toxic effect." EPA's guidelines for assessing potential joint toxic actions first call for a preliminary grouping of those chemicals that might cause a common toxic effect by a common mechanism. Criteria used to preliminarily group chemicals include: structural similarity; general mechanism of mammalian toxicity; and a particular toxic effect (EPA, 1999e). EPA (1999e) emphasizes that chemicals identified under this first step "will not be included in a cumulative risk assessment if it is determined that they do not cause a common toxic effect by a common mechanism." In step 2, EPA (1999e) definitively identifies those chemicals that cause a common toxic effect and, in step 3, determines the mechanism by which each chemical causes that effect. In the remaining steps, EPA (1999e) compares each mechanism to identify those chemicals with both a common toxic effect and a common mechanism. By following this multiple-step approach that involves a thorough

evaluation of toxicity data, EPA will determine that many chemicals identified in the preliminary step will not be carried forward.

Finally, it must be noted that the arithmetic summing of cancer risk is incorrect for two reasons. First, it is not appropriate to sum cancer risks of chemicals with different cancer classifications because the human cancer risk of those chemicals which are less likely to be human carcinogens will be overstated. As explained by DOE (1995), the addition of slope factors for multiple chemicals in a mixture sums all carcinogens equally, regardless of their carcinogenicity classification. Adding cancer risk is also mathematically incorrect. Cancer risks are typically based on cancer slope factors that represent the upper 95th percentile estimates of potency. These 95th percentile estimates are not strictly additive (EPA, 1989; Lang, 1995). Burmaster and Lehr (1991) show that when three 95th percentile values are combined, the outcome represents the 99.99th percentile. EPA (1989) agrees, noting that "when adding cancer slope factors from multiple chemicals, the total cancer risk estimate might become more conservative."

EPA has also asked that GTA's risk assessment apply the additivity approach to non-carcinogens. However, the Agency has cautioned against this approach:

> Application of the hazard index equation to a number of compounds that are not expected to induce the same type of effects or that do not act by the same mechanism, although appropriate as a screening-level approach, could overestimate the potential for effects. This possibility is generally not of concern if only one or two substances are responsible for driving the HI above unity. If the HI is greater than unity as a consequence of summing several hazard quotients of similar value, it would be appropriate to segregate the compounds by effect and by mechanism of action and to derive separate hazard indices for each group.

EPA (1989)

Use of the "Toxicity Equivalency Quotient" Approach

In the draft dioxin reassessment, EPA has proposed to use the Toxicity Equivalency Quotient (TEQ) approach to assess the toxicity of PCBs. As discussed in detail in Appendix 3, the TEQ approach, as applied to PCB congeners, is based largely on assumptions that are of questionable scientific validity and are inconsistent with existing experimental data. The result is to inappropriately and substantially overpredict both the cancer and non-cancer toxicity of PCBs.

The TEQ approach is based on the finding that certain chemicals, including 12 coplanar PCB congeners, exhibit structural and toxicokinetic similarities to TCDD. These chemicals are said to act through the same mechanism as TCDD and related compounds, namely binding with the aryl hydrocarbon receptor (AhR). Other chemicals that bind with the AhR include polynuclear aromatic hydrocarbons, certain hormones, certain drugs, and a chemical formed by the human metabolism of the indole glucobrassicin (which is found in cruciferous vegetables, including cabbage, cauliflower, and broccoli) (Fiala et al., 1985; Hodgson and Levi, 1987). The flaws of the TEQ approach, as applied to PCB congeners, can be summarized as follows:

- The TEF approach is inappropriate for assessing PCB toxicity because it is based on the unproven assumption that PCBs have additive toxic effects. The fact that a chemical binds with the AhR does not mean that it will cause an adverse effect. In fact, chemicals that bind with the AhR can have a beneficial effect (e.g., triggering a normal physiological response like enzyme induction), an adverse effect, or no effect.
- A significant amount of laboratory data indicates that TEFs are not, in fact, additive. Moreover, the TEQ approach ignores the long-established approach that evaluates PCB toxicity using cancer slope factors (CSFs) for PCB mixtures. If the TEQ methodology for PCBs is correct, then it should accurately predict the toxicity of PCB mixtures. Appendix 3 demonstrates that the TEQ approach overpredicts the carcinogenicity of PCB mixtures by over an order of magnitude.
- Appendix 3 also explains how laboratory evidence demonstrates that the TEQ approach over-predicts the noncancer toxicity of PCBs. For example, Bannister et al. (1987) treated mice with TCDD alone and with a mixture of TCDD and between 1,300 and 20,000 times as much Aroclor 1254. The mice treated with TCDD alone showed a large depression in the formation of certain infection fighting cells. The mice treated with the TCDD/Aroclor 1254 mixture showed no depression. Apparently, the Aroclor 1254, which contains agonists, partial agonists and antagonists, overall had an antagonist effect that wholly offset the agonist effect of the TCDD.

  EPA's advocacy of the unproven TEQ approach is a good example of Agency action that is not based on sound science and, moreover, is demonstrably poor at accurately estimating toxicity. It should be an high priority for EPA to reject scientific approaches that do not comport with the evidence.

II. Ecological Risk Examples

A. Toxicity Assessment

Ecological risk assessments are performed by comparing exposure to Toxicity Reference Values (TRVs), which are analogous to the RfDs used in human health risk assessments. There is currently no standard approach for assessing potential carcinogenic effects in ERAs. Recent guidance, the Ecological Soil Screening Level Guidance or EcoSSL (EPA, 2003b), summarizes four methods to derive TRVs: (1) critical study approach; (2) benchmark dose model approach; (3) distribution approach; and (4) weight-of-evidence approach. These are summarized below:

- The critical study approach has been the most common method used to derive TRVs (e.g., Sample et al., 1996). Similar to the derivation of a human health RfD, uncertainty factors are applied to a LOAEL or NOAEL from the critical study. The selection of the appropriate critical study and uncertainty factor(s) are key to the determination of whether the derived TRV is overly conservative or appropriate for the particular site.
- The benchmark dose approach is also based on a critical study, but uses the benchmark dose model (EPA, 2000), in lieu of NOAELs or LOAELs, to derive the TRV. This approach does not require uncertainty factors, but instead uses the entire dose-response relationship from the critical study and fits an appropriate curve to these values. The TRV is then assigned as the value that represents an incremental effect of 10% at the 95% confidence level. The selection of the appropriate critical study is key to the determination of whether the TRV is overly conservative or appropriate for the particular site. Although not commonly employed by USEPA for TRV development, the BMD approach has been used by other government entities (e.g., USACHPPM, 2000).
- The distribution approach uses probability density function that represents the sensitivity of different species to the evaluated endpoint at a particular level (such as the LD50). Although this approach has the potential to address the slopes of the dose-response curves, it suffers from the lack of an adequate toxicological database to best define the probability density function.
- The weight-of-evidence approach combines the results from a number of studies using different measurement endpoints (e.g., growth, reproduction) and test organisms. The TRV is calculated as the geometric mean of the NOAELs for growth and reproduction effects, since these are most relevant to potential population effects.

EPA Recommended "Weight-of-Evidence" Approach to Derive TRVs

Although the EcoSSL Guidance discusses the four approaches to derive TRVs (EPA, 2003b), the Agency chose to select the weight-of-evidence approach to calculate SSLs because the weight-of-evidence approach considers:

> all of the extracted toxicological data in place of the selection of one critical study. The use of the critical study approach would require considerable professional judgment thereby decreasing the transparency and reproducibility of the wildlife TRV derivation process. To avoid foreseen conflicts over selection of "one" result; to prevent the need for "committee" selection and to attain transparency and reproducibility this method [critical study approach] was not selected.

Although the weight-of-evidence approach may be appropriate for deriving screening level values, it should not be used for site-specific ecological risk assessments because the weight-of-evidence approach does not address the measurement endpoints that may be relevant to key ecological receptors of interest for a particular site. For example, when evaluating the studies considered for the weight-of-evidence assessment, less weight should be given to studies that evaluate the toxicity of the given chemical to receptors that are not found at the particular site, to studies of chemicals whose form may not be relevant to the site-specific form (e.g., use of lead salts when lead may be bound to sulfides in the environment), and to studies whose designs are inconsistent with the exposure that may occur under the environmental conditions at issue (e.g., plant toxicity studies based on hydroponic exposure to the salt form of the metal). EPA (2003b) did not compare the weight-of-evidence approach to the more commonly applied critical study approach, or to the other two methods (the benchmark dose or distribution approaches), implying that the weight-of-evidence approach may not have received adequate peer review. EPA has not demonstrated that the weight-of-evidence approach will result in TRV values that are appropriate for EcoSSL development, much less for evaluation of site-specific conditions.

The goal of the EcoSSl development process is to provide levels that can be used to screen chemicals on a site-specific basis. The TRVs developed for EcoSSLs are based on no effect levels and do not represent values suitable for the protection of populations that have been suggested as assessment endpoints. Exceedance of the TRV does not necessarily imply that there will be an effect, much less an adverse effect. Thus, the use of these TRVs will result in EcoSSLs that will be so low as to not exclude (screen out) any compounds.

EPA Avian TRV for Dioxin for the Hudson River

USEPA has inconsistently applied uncertainty factors for deriving the TRV for dioxins in avian species based on the Nosek et al. (1992) study, which involved a 10-week exposure period. Oak Ridge National Lab's assessment of this study concluded that effects on survival, egg production, and egg hatchability were observed only at the maximum dose (Sample et al., 1996). Because this study considered exposure throughout a critical life stage (reproduction), the upper no effect dose level (0.1 µg/kg-week) was considered to be a chronic NOAEL. Therefore, an uncertainty factor was not applied, or required, to derive the TRV, which was calculated at 1.4 x 10^{-5} mg/kg-day. USEPA's Screening Level Ecological Risk Assessment Protocol for Combustion Facilities, simply rounded the TRV to 1 x 10-5 mg/kg-day, but used the same approach (EPA, 1999f). USEPA's Data Collection for the Hazardous Waste Identification Rule, (Section 14: Ecological Benchmarks) took a slightly different approach, assigning the TRV as the geometric mean of the NOAEL and LOAEL from the same study (4.4 x 10^{-5} mg/kg-day), but did not apply an uncertainty factor (EPA, 1999g). Despite the precedent established by its own documents, EPA (2000b), in its ecological assessment of the Upper Hudson River, used a TRV of 1.4 x 10^{-6} mg/kg-day for the evaluation of dioxin-like PCB congeners. EPA (2000b) assumed that the 10-week exposure period represented a subchronic rather than chronic exposure and applied an uncertainty factor of 10 to account for this, resulting in a 10-fold more conservative TRV than used elsewhere.

EPA Otter TRV for PCBs for the Hudson River

The Upper Hudson River Revised Baseline Ecological Risk Assessment (RBERA) developed conservative TRVs. The RBERA did not report any species-specific studies for the development of the TRVs for River otters. Instead, the NOAEL (0.004 mg/kg-day) and LOAEL (0.04 mg/kg-day) identified for mink were used for the phylogentically similar River otter (both are members of the Family Mustelidae). The study used to develop the TRVs for the mink (Restum et al., 1998) included confounding exposures to pesticides. The authors did not attempt to segregate the potential contribution of the pesticides to the evaluated endpoint (kit survival), nor was this uncertainty included in the TRV development. The derivation of appropriate TRVs is critical not only in the RBERA, but also in the assessment of remedial measures. For example, when the TRVs for mink and River otter were adjusted to reflect appropriate interspecies relationships and realistic exposures (e.g., area use factors), the source control alternative achieves lower risks to mink and otter than dredging in 34 of the 40 miles of the Upper Hudson River.

B. Risk Characterization

Risk characterization in an ERA includes several components (EPA, 1997a): (1) the integration of exposure profiles with exposure-effects information; (2) calculation of hazard quotients (HQs) or hazard indices (HIs); (3) discussion of approaches to interpret the risk results; (4) identification of a threshold for adverse effects on the assessment endpoints; and (5) summary of the associated uncertainties. The most common metrics for ecological risks are the HQ and HI. An HQ less than one (unity) indicates that the contaminant alone is unlikely to cause adverse ecological effects. Similarly, when the HI (the sum of HQs for chemicals with similar mechanisms of toxicity and assessment endpoints) is less than one, the group of chemicals is unlikely to result in adverse ecological effects.

Risks are then combined across exposure pathways for the representative receptor(s) to develop receptor group-specific risks/hazards. The focus is typically on water ingestion and dietary exposure routes, although this can be species-specific (e.g., dermal transport through feet in wading birds; inhalation of volatiles by burrowing animals). The uncertainty in the risk estimate is then assessed. Although uncertainty assessment is typically qualitative, quantitative approaches are preferable, especially if probabilistic methods are used as part of the uncertainty assessment.

EPA does not provide any formal guidance on the determination of a "level of significance" when the HQ or HI is above one. EPA (1997a) does recommend that:

> The lower bound of the threshold would be based on consistent conservative assumptions and NOAEL toxicity values. The upper bound would be based on observed impacts or predictions that ecological impacts could occur. This upper bound would be developed using consistent assumptions, site-specific data, LOAEL toxicity values, or an impact evaluation.

Notwithstanding this guidance, the use of a bounding approach is rarely seen in ERAs.

One approach to assess the significance of an exceedance of the HI or HQ is in the context of potential population-level effects. EPA (1997a) recommends that potential ecological risks should be assessed at the population-level for all but threatened and endangered species. Although no explicit guidance is provided, this is typically accomplished through the use of measurement endpoints that are related to population effects (e.g., using TRVs based on growth or reproductive effects).

Fox River ERA

The Fox River/Green Bay ROD relied, in part, on an ERA to justify selection of a dredging remedy. However, the ERA is seriously defective and inconsistent with USEPA guidance. In fact, the ERA is little more than a screening level assessment which, under EPA guidance, may not be used as the basis of a remedy decision.

Under EPA guidance (EPA, 1997a), an ERA is to be performed as a stepwise process. This process moves from a conservative screening analysis to definitive "baseline" risk characterization, with the latter employing site-specific data as much as possible. In short, screening level risk assessments are conducted to support a fundamental threshold decision: Does a site require additional risk assessment? Screening level assessments are not conducted to support major risk management decisions.

The Fox River ERA is inconsistent with this guidance. Although the ERA discusses the substantial site-specific data base which was compiled by the Fox River PRPs, the US Fish and Wildlife Service, USEPA, WDNR, universities, and other organizations and institutions, the ERA utterly fails to use this information in the risk assessment. Instead, the ERA is based on conservative assumptions and generic exposure scenarios, and demonstrably overstates ecological risk posed by the presence of PCBs in the Fox River. A few examples of the ERA's deficiencies follow:

- The ERA is highly misleading in its treatment of site-specific habitat data. In response to industry comments, the ERA summarizes habitat information for a number of ecological receptors and cites a detailed habitat analyses conducted by the Fox River PRPs. However, the ERA then proceeds to include the site-specific habitat information only in qualitative discussions in an introductory section of the ERA. The site-specific habitat information is not used for the purpose of risk quantification. The ERA offers no explanation for this omission.
- Ignoring site-specific habitat information leads the ERA to false conclusions. For example, the ERA's highest hazard quotient projections for mink are in one reach of the Fox River and in one area of Green Bay. However, data cited by the ERA clearly show that there are very few mink in these areas. Thus there is, in fact, little or no risk to mink populations.
- The ERA erroneously assumes that risk derives equally from PCBs distributed in all areas of the river. Thus, the ERA calculates a single sediment quality threshold to be applied to all sediments whether or not they contribute substantially to fish exposure. If the ERA had accurately accounted for fish habitat preferences, it would have been clear that a single reach-wide sediment quality threshold is inappropriate and not scientifically supported.

By ignoring readily available site-specific information, the ERA provided inaccurate risk characterization, an ineffective foundation for risk management decisionmaking, and a risk assessment that is not in keeping with the provisions of applicable USEPA guidance.

Upper Hudson River ERA

The Upper Hudson River Revised Baseline Ecological Risk Assessment (RBERA) was considered in making the remedial decision for the Upper Hudson River in New York. Under EPA's own guidance (EPA, 1997a), however, the results of the RBERA should not have been used to determine remedial action, because the approach actually employed by EPA's contractor was designed for screening-level applications. This RBERA is based principally on conservative data and assumptions that are deliberately designed to be conservative to minimize the possibility that any potential adverse effects will be missed in a screening-level analysis. As such, these data and assumptions overstate the actual effects of most chemicals at most sites. For the baseline ecological risk assessment of so prominent a site as the Upper Hudson River, EPA should have refined its toxicity quotient-based approach to incorporate more site-specific information. In addition, EPA should have used an approach that incorporates data on the actual conditions of fish and wildlife populations in and along the Hudson River. In fact, EPA (2000d) discussed several field studies in the RBERA, but dismissed their relevance and did not integrate the results into the ecological risk assessment.

On behalf of EPA, Eastern Research Group coordinated a review of the Upper Hudson River ecological risk assessment by seven independent peer reviewers. This peer review group sharply criticized EPA's work product, concluding that EPA's draft ecological risk assessment represented a screening-level effort. The peer reviewers provided EPA with specific recommendations to reduce the conservatism and recommended that more sophisticated approaches be used for evaluating ecological risks. For the most part, EPA either failed to implement these recommendations, implemented the recommendations incorrectly, or made offsetting changes to the recommendations that resulted in little reduction to the level of conservatism.

For example, the peer reviewers found that EPA did not embrace an appropriate weight-of evidence approach in conducting the risk assessment. The peer reviewers "questioned why EPA's conceptual site model artificially constrains the risk assessment to the main channel of the Upper Hudson River, given the fact that many receptors (e.g., birds, mammals, and fish) may use a far broader range of habitat," and "reviewers were concerned that the risk assessment, with its current spatial construct, becomes too narrow in scope" (U.S. EPA 2000d, p. 2-10). In conclusion, the peer review group unanimously agreed that EPA's characterization of the ecological setting was inadequate: Without a description of the habitats, the species occupying the Upper Hudson River, and the spatial and temporal use of habitats by species considered in the conceptual site model, the reviewers did not think it was possible to defend the risk characterization" (U.S. EPA 2000d, p. 2-2).

PCB Worm Tissue Criterion for the Historic Area Remediation Site

In October 2002, EPA developed Proposed Polychlorinated Biphenyl Worm Tissue Criterion for the Historic Area Remediation Site (HARS) (67 FR 62659). This document established a HARS-specific worm tissue PCB criterion of 113 parts per billion (ppb) for use in determining the suitability of proposed dredged material for use as remediation material. The 113 ppb criterion is based on a number of conservative assumptions including the following: (1) 100% of fish consumed by New Jersey anglers are sport-caught saltwater finfish, even though the data allow one to distinguish between salt and fresh water fishing; (2) 100% of the fish consumed are caught at the HARS; (3) all species consumed by recreational anglers are available at the HARS; (4) anglers fish consistently every year for 70 years; (5) there is no loss of contaminants due to cooking methods; (6) the site use factor of 77.7% for all fish species is not supported by commercial landings in the vicinity of HARS; and (7) use of a trophic transfer factor of 3 for all organics does not properly capture their potential for bioaccumulation. The final rule was issued on March 17, 2003 (68 FR 12592). The criterion remained 113 ppb with no adjustment to the conservative assumptions.

Tier II Great Lakes Initiative Water Quality Criteria

In the Great Lakes Water Quality Initiative (GLWQI), EPA proposed a two-tiered approach to deriving water quality criteria. A Tier I water quality criterion is derived when specific data requirements are met (e.g., sufficient toxicological data exist for eight taxonomic groups). These data requirements are identical to those that EPA has used historically as the minimum requirements for calculation of ambient water quality criteria. Under the GLWQI regulations, a Tier II water quality value can be derived if the data required to derive a Tier I value are not available, or if the data are not of high quality.

Because Tier II criteria are to be derived based on incomplete or inferior data, EPA builds in several levels of conservatism in the calculations. However, as discussed below, it is easily seen that using the EPA's Tier II derivation process often yields grossly over-conservative water quality criteria.

The approach used to derive acute and chronic Tier II values is as follows[21]:

- The available toxicity test results are ordered by genus and the geometric mean of the test results is calculated for each genus. The genus mean values are ranked from lowest to highest and counted.
- The lowest genus mean value is divided by the secondary acute factor (SAF), which is effectively a safety factor. The SAF ranges from 21.9 to 4.3 and decreases as the number of suitable studies available increases. That is, the more data available, the smaller the safety factor. The resulting number is called a secondary acute value (SAV).
- The Tier II acute value is the SAV divided by two.
- The secondary acute-to-chronic ratio (SACR) is the geometric mean of the ratios of acute and chronic toxic concentrations from at least three studies that have investigated both effects. If less than three studies are available, the missing values are replaced with a default value of eighteen.
- The Tier II chronic value is the SAV divided by the SACR.

The approach used to derive Tier II values can result in extremely low values, particularly when only a few acceptable toxicity studies are available (Alsop and Unwin, 1994). This is because the amount of conservatism in the Tier II value increases as the number of suitable studies decreases. For example, the comparison of chronic Tier I values for nine metals to their corresponding Tier II values show that the Tier II values overestimate the Tier I values from 3 to 16,000 times at the 95th percentile of the secondary acute factor (Alsop and Unwin, 1994). As another example, Suter and Tsao (1996) used the Tier II approach to develop potential screening benchmarks for protection of aquatic life from common contaminants in water. Because the Tier II values for sodium chloride were below commonly occurring ambient concentrations of this

[21] The example presented assumes that toxicity data for a daphnid are available. If no toxicity data for a daphnid are available, Tier II values cannot be developed.

salt, they were judged to be inappropriate by the study authors and were not presented (Suter and Tsao, 1996).

* * *

III. References

ACGIH. 1993. Threshold limit values (TLVs) for chemical substances in the work environment for 1993-1994. American Conference of Government Industrial Hygienists. Cincinnati, OH.

Alsop, W.R. and J.Unwin. 1994. Tier II Values for the Protection of Freshwater Aquatic Life. Proceedings of the 1994 TAPPI Conference. Portland, Oregon.

AMEC Earth & Environmental. 2002. Development of a Revised Chronic Reference Dose for Polychlorinated Biphenyls (Aroclor 1254) Based on Empirical Data.

AMEC Earth & Environmental. 2002a. Alternative Human Health Risk Assessment of the Lower Fox River and Green Bay, Wisconsin. January 18.

American Industrial Health Council 1999. Letter to William P. Wood, USEPA Risk Assessment Forum, on Revised EPA Guidelines for Carcinogen Risk Assessment.

Argus Research Laboratories inc (1999) Oral (drinking water) two-generation (one litter per generation) reproduction study of ammonium perchlorate in rats. Horsham, PA: Argus Research Laboratories, Inc; protocol no. 1416-001.

Argus Research Laboratories, Inc. (2001) Hormone, thyroid and neurohistological effects of oral (drinking water) exposure to ammonium perchlorate in pregnant and lactating rats and in fetuses and nursing pups exposed to ammonium perchlorate during gestation or via maternal milk. Horsham, PA: protocol no. ARGUS 1416-003

ATSDR. 1995. Toxicological Profile for Naphthalene, 1-Methylnapthalene, 2-Methylnapthalene. Update. August.

Bannister, R., D. Davis, T. Zacharewski et al. 1987. Aroclor 1254 as a 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin antagonist: Effects on enzyme induction and immunotoxicity. Toxicol. 46:29-42.

Barton, H.A. and H. J. Clewell III. 2000. Evaluating noncancer effects of trichloroethylene: Dosimetry, mode of action and risk assessment. Environmental Health Perspectives, Supplement 2 108:323-334.

Beck B.D. Connolly R.B., Dourson M.L., Guth D., Hattis D., Kimmel C. and Lewis S.C. 1992. Improvements in Quantitative Noncancer Risk Assessment Fundamental and Applied Tox. 20:1-14.

Bekkedal MYV, Carpenter T, Smith J, Ademujohn C, Maken D, Mattie DR (2000) A neurodevelopmental study of the effects of ammonium perchlorate on the motor activity of pre-weanling rat pups. Wright Patterson Air Force Base, OH; Naval Health Research Center Detachment, Neurobehavioral effects laboratory; report no. TOXDET- 00-03

Berman, E., M. Schlicht, V. Moser, R. MacPhail. 1995. A multidisciplinary approach to toxicological screening. I: Systemic toxicity. Journal of Toxicology and Environmental Health 45:127-143.

Brechner RJ, Parkhurst GD, Humble WO, Brown MB, Herman WH. 2000. Ammonium perchlorate contamination of Colorado River drinking water is associated with abnormal thyroid function in newborns in Arizona. J Occup Environ Med. 42(8):777-82

Buben, J.A. and E.J. O'Flaherty. 1985. Delineation of the role of metabolism in the hepatotoxicity of trichloroethylene and perchloroethylene: a dose-effect study. Toxicol Appl. Pharmacol. 78:105-122.

Buckpitt, A., Boland, B., Isbell, M., Morin, D. Shultz, M., Baldwin, R., Chan, K., Karlsson, A., Lin, C., Taff, A., West, J., Fanucchi, M., Van Winkle, L., and Plopper, C. 2002. Naphthalene-induced respiratory tract toxicity: Metabolic mechanisms of toxicity. Drug Metabolism Reviews. 34(4):791-820.

Burmaster, D.E. and J.H. Lehr. 1991. It's Time to Make Risk Assessment a Science. Ground Water Monitoring Review XI(3):5-15.

Calabrese, E.J., E.J. Stanek, C.E. Gilbert, and R.M. Barnes. 1990. Preliminary adult soil ingestion estimates: results of a pilot study. Regul Toxicol Pharmacol. Aug;12(1):88-95.

Calabrese, E.J., E.J. Stanek, P. Pekow, and R.M. Barnes. 1997. Soil ingestion estimates for children residing on a superfund site. Ecotoxicol Environ Saf. Apr;36(3):258-68.

ChemRisk. 1991. Consumption of Freshwater Fish from Maine Lakes and Ponds. Portland, Maine. September 6.

Cicmanec, J.L. and K.A. Poirier. 1995. Selected applications of reduced uncertainty factors in noncancer risk assessment. Hum. Ecol. Risk Assess. 1(5):637-640.

Clewell, H.J. III, P. Robinan Gentry, T.R. Covington, and J.M. Gearhart. 2000. Development of a physiologically based pharmacokinetic model of trichloroethylene and its metabolites for use in risk assessment. Environmental Health Perspectives, Supplement 2 108:283-306.

Connelly, N.A., B.A.Knuth & C.A.Bisogni. 1992. Effects of the Health Advisory Changes on Fishing Habits and Fish Consumption in New York Sport Fisheries. Human Dimension Research Unit. Cornell University (Report for the New York Sea Grant Institute Project No. R/FHD-2-PD).

Connelly, N.A., B.A. Knuth, and T.L. Brown. 1996. Sportfish Consumption Patterns of Lake Ontario Anglers and the Relationship to Health Advisories. North American Journal of Fisheries Management 16:90-101.

Crump K. 1999. Benchmark doses for perchlorate obtained from Lamm et al (1999) study of thyroid function in perchlorate workers. Comments submitted to EPA external peer review panel.

Crump C, Michaud P, Tellez R, Reyes C, Gonzalez G, Montgomery EL, Crump KS, Lobo G, Becerra C, Gibbs JP. 2000. Does perchlorate in drinking water affect thyroid function in newborns or school-age children? J Occup Environ Med. 42(6):603-12

Crump C, Weiss NS and Goodman G. 2001. Methods and conclusions of the Arizona Perchlorate Study. J Occup Environ Med. 43(4):307-9.

Crump K, Goodman G. 2003. Benchmark analysis for the perchlorate inhibition of thyroidal radioiodine uptake utilizing a model for the observed dependence of uptake and inhibition on iodine excretion. Comments submitted to Cal/EPA OEHHA

Delzell, E; N. Sathiakumar, and M. Hovinga. 1996. A follow-up study of synthetic rubber workers. Toxicology 113:182-189.

DOE. Department of Energy. 1995. CERCLA Baseline Risk Assessment Reference Manual for Toxicity & Exposure Assessment and Risk Characterization. U.S. Department of Energy, Office of Environmental Policy & Assistance RCRA/CERCLA Division. Washington, D.C. March.

Dourson, M.L. and J. Stara. 1988. Regulatory history and experimental support of uncertainty (safety) factors. Regulatory Toxicology and Pharmacology 3:224-238.

Dourson M.L., Felter S.P. and Robinson D. 1996. Evolution of Science-Based Uncertainty Factors in Noncancer Risk Assessment. Reg. Tox. and Pharmacology 24: 108-120.

Ebert, E.S., N.W. Harrington, K.J. Boyle, J.W. Knight, and R.E. Keenan. 1993. Estimating consumption of freshwater fish among Maine anglers. N. Am. J. Fish. Mgt. 13:737-745.

Ebert, E.S., S.H. Su, T.J. Barry, M.N. Gray, and N.W. Harrington. 1996. Estimated Rates of fish consumption by anglers participating in the Connecticut Housatonic River creel survey. *N.* Am. J. Fish. Mngt. 16:81-89.

EPA. 1986. Guidelines for health risk assessment of chemical mixtures. Federal Register 51(185):34014-34025.

EPA. 1986a. Guidelines for Carcinogen Risk Assessment. Federal Register 51(185):33992-34003.

EPA. 1989. Risk Assessment Guidance for Superfund; Volume I: Human Health Evaluation Manual (Part A) - Interim Final. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C. EPA/540/1-89/002. December.

EPA. 1991. RAGS Volume I: Human Health Evaluation Manual, Supplemental Guidance: Standard Default Exposure Factors, OSWER Directive 9285.6-03.

EPA. 1992. Provisional Non-cancer and Cancer Toxicity Values for Potassium Perchlorate (CASRN 7778-74-7) (Aerojet General Corp./CA), Memorandum from Joan S. Dollarhide, Superfund Health Risk Technical Support Center, Environmental Criteria and Assessment Office, Office of Research and Development, to Dan Stralka, US EPA Region IX.

EPA. 1992a. Guidelines for Exposure Assessment. Federal Register 57(104): 22888-22938.

EPA. 1993. Use of IRIS Values in Superfund Risk Assessment.. OSWER Directive 9285.7-16.

EPA. 1995. Correspondence from Joan S. Dollarhide, National Center for Environmental Assessment, Office of Research and Development, to Mike Girrard, Chairman, Perchlorate Study Group.

EPA. 1995a. Land Use in the CERCLA Remedy Selection Process. OSWER Directive No. 9355.7-04. May 25.

EPA. 1995b. Risk Characterization memorandum from Administrator Carol Browner

EPA. 1996. Proposed Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Office of Research and Development. EPA/600/P-96/003C. April.

EPA. 1996a. PCBs: Cancer Dose Assessment and Application to Environmental Mixtures. U.S. Environmental Protection Agency, Office of Research and Development. EPA/600/P-96/001. Washington, D.C.

EPA. 1997a. Ecological Risk Assessment Guidelines for Superfund: Process for Designing and Conducting Ecological Risk Assessments, Interim Final. Office of Solid Waste and Emergency Response. EPA 540-R-97-006. June 1997.

EPA. 1997b. Exposure Factors Handbook, Volumes 1, II and III. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Washington, D.C. EPA/600/P-95/002Fa,b,c. August.

EPA. 1997c. EPA Policy for Use of Probabilistic Analysis in Risk Assessment

EPA. 1998a. Guidelines for neurotoxicity risk assessment. Federal Register 63(93):26926-26954.

EPA. 1998b. Health risk assessment for 1,3-butadiene, external review draft. U.S. Environmental Protection Agency, Office of Research and Development. EPA/600/P-98/001A. January.

EPA. 1999. Guidance for Conducting Health Risk Assessment of Chemical Mixtures. U.S. Environmental Protection Agency. NCEA-C-0148. External Scientific Peer Review Draft. April. http://www.epa.gov/nceawww1/pdfs/mixtures.pdf

EPA. 1999a. Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites. Office of Solid Waste and Emergency Response. OSWER Directive 9285.7-28. October.

EPA. 1999b. Draft Guidelines for Carcinogen Risk Assessment. July.

EPA. 1999c. "Risk Assessment Guidance for Superfund Volume I - Human Health Evaluation Manual Supplemental Guidance Dermal Risk Assessment Interim Guidance." Office of Emergency and Remedial Response, Washington, DC. March 1999. Draft.

EPA. 1999d. Memorandum from A. Koppikar, National Center for Environmental Assessment, Office of Research and Development, to L. McKelvey, Emissions Standards Division, Office of Air Quality Plannnig and Standards. April 6.

EPA. 1999e. Guidance for Identifying Pesticide Chemicals and Other Substances that have a Common Mechanism of Toxicity. U.S. Environmental Protection Agency. January.

EPA, 1999f. Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities. OSWER. EPA530-D-99-001A. August.

EPA, 1999g. Data Collection for the Hazardous Waste Identification Rule, Section 14: Ecological Benchmarks. Draft. October.

EPA. 2000. Benchmark Dose Technical Guidance Document. EPA/630/R-00/001. October. 87p.

EPA. 2000a. Presenter's Manual for: "Superfund Risk Assessment and How You Can Help." EPA/540/R-99/013. Office of Solid Waste and Emergency Response. July 2000.

EPA. 2000b. Revised baseline ecological risk assessment, Hudson River PCBs reassessment RI/FS. U.S. Environmental Protection Agency.

EPA. 2000c. Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, D.C. EPA/630/R-00/002. August.

EPA. 2000d. Report on the peer review of the ecological risk assessment for the Hudson River PCBs Superfund Site. Final report. Prepared by Eastern Research Group, Inc

EPA. 2000f. Further Site Characterization and Analysis, Revised Human Health Risk Assessment (HHRA) Volume 2F, Hudson River PCBs Reassessment RI/FS. Prepared for the USEPA Region 2 and USACE by TAMS Consultants, Inc. and Gradient Corporation.

EPA. 2001. Dioxin Reassessment – An SAB Review of the Office of Research and Development's Reassessment of Dioxin. U.S. Environmental Protection Agency, Science Advisory Board, Washington, D.C. EPA-SAB-EC-01-006. May.

EPA. 2001a. Risk Assessment Guidance for Superfund: Volume 3 - (Part A, Process for Conducting Probabilistic Risk Assessment). Final. United States Environmental Protection Agency, Solid Waste and Emergency Response. EPA 540-R-02-002. December.

EPA. 2001b. Planning for Ecological Risk Assessment: Developing Management Objectives. EPA/630/R-01/001A. June.

EPA. 2001c. Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites. Peer Review Draft. OSWER 9355.4-24. March 2001.

EPA. 2001d. Trichloroethylene Health Risk Assessment: Synthesis and Characterization. Office of Research and Development, US Environmental Protection Agency. EPA/600/P-01/002A. August.

EPA. 2001e. A Review of the Reference Dose and Reference Concentration Processes. Reference Dose (RfD) Technical Panel, Risk Assessment Forum, US Environmental Protection Agency. NCEA-F-1104. August. Internal Review Draft.

EPA. 2001f. Risk Assessment Guidance for Superfund Volume I: Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment) Interim Review Draft. Office of Emergency and Remedial Response. U.S. Environmental Protection Agency. Washington, DC. EPA/540/R/99/005. September.

EPA. 2002. Preliminary Remedial Goals guidance. Region 9.

EPA. 2002a. An SAB Report: A Framework for Assessing and Reporting on Ecological Condition. Science Advisory Board. EPA-SAB-EPEC-02-xxx. March. External Review Draft .

EPA. 2002b. Region 9 PRG Table 2002.
www.epa.gov/region09/waste/sfund/prg/files/02userguide.pdf

EPA. 2002c. Health assessment document for 1,3-butadiene. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. EPA/600/P-98/001. October.

EPA. 2002d. Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization. EPA External Review Draft (Jan. 16, 2002).

EPA. 2003a. Draft Final Guidelines for Carcinogen Risk Assessment. External Review Draft. February.

EPA. 2003b. Ecological Soil Screening Levels (Eco-SSLs) Standard Operating Procedure (SOP) #4: Derivation of Wildlife Toxicity Reference Value (TRV). Attachment 4-5, Guidance for Developing Ecological Soil Screening Levels. March 2003.

Faustman, E. 1996. Review of Noncancer Risk Assessment: Application of Benchmark Dose Methods. Prepared for the Commission on Risk Assessment and Risk Management.

Fawer, R.F., U. DeRibaupierre, M.P. Guillemin, M. Berode and M. Lobe. 1983. Measurement of hand tremor induced by industrial exposure to metallic mercury. J. Ind. Med. 40: 204-208.

Fiala, E.S., B.S. Reddy, and H.H. Weisburger. 1985. Naturally occurring anticarcinogenic substances in foodstuffs. In Annual Review of Nutrition, pp. 295-321. Volume 5. Palo Alto: Annual Review Inc.

Frederick, CR; Udinsky, JR; Finch, L. 1994. The regional hydrolysis of ethyl acrylate to acrylic acid in the rat nasal cavity. Toxicol Lett 70:49-56.

Gibbs JP, Ahmad R, Crump KS, Houck DP, Leveille TS, Findley JE, Francis M. 1998. Evaluation of a population with occupational exposure to airborne ammonium perchlorate for possible acute or chronic effects on thyroid function. J Occup Environ Med. 40(12):1072-82

Greer MA, Goodman G, Pleus RC, Greer SE. 2002. Health effects assessment for environmental perchlorate contamination: the dose response for inhibition of thyroidal radioiodine uptake in humans. Environ Health Perspect. 110(9):927-37

Gross, E.A., Swenberg, J.A., Fields, S., and Popp, J.A. 1982. Comparative morphometry of the nasal cavity in rats and mice. J. Anat. 135, 83-88.

Hawley, J.K. 1985. Assessment of health risk from exposure to contaminated soil. Risk Analysis 5(4):289-302.

Harrington, N.W., Curry, C.L., and P.S. Price. 1995. The MicroExposure® Event Modeling Approach to Probabilistic Exposure Assessment. Paper No. 95-TA42.03. Proceedings of the 88th Annual Meeting of the Air and Waste Management Association. San Antonio, Texas, USA. June.

Hodgson, Ernest and P.E. Levi. 1987. Textbook of Modern Toxicology. Toxicology Program, Northern Carolina State University. New York: Elsevier Publishing.

Holmes, K.K., J.C. Kissel, and K.Y. Richter. 1996. (prepublication draft, cited in EPA, 1997)

ICH. (International Conference on Harmonisation of Technical Requirements For Registration of Pharmaceuticals for Human Use). 1998. Duration of Chronic Toxicity Testing in Animals (Rodent and Non Rodent Toxicity Testing). Recommended for Adoption at Step 4 of the ICH Process on 2 September 1998 by the ICH Steering Committee.

IRIS. 2003. Integrated Risk Information System. www.epa.gov/iris/

Keenan, R.E., P.S. Price, J. McCrodden, and E.S. Ebert. 1996. Using a microexposure event analyses to model potential exposures to PCBs through ingestion of fish from the Upper Hudson River. In: Organohalogen Compounds: Proceedings Dioxin '96-16th International Symposium on Chlorinated Dioxins and Related Compounds, Amsterdam, The Netherlands. *Organohalogen* 30:61-65.

Kelsh M, Buffler P, Daaboul J, Rutherford G, Lau Edmund, Exuzides A, Cahill J, Palmer L, Lorey F. 2002. Primary congenital hypothyroidism and newborn thyroid function among residents of a southern California community. Submitted to JOEM and as comments to Cal EPA OEHHA.

Kissel, J.C., K.Y. Richter, and R.A. Fenske. 1996. Field measurement of dermal soil lading attributable to various activities: Implications for exposure assessment. Risk Analysis 16(10):115-125.

Lamm SH, Braverman LE, Li FX, Richman K, Pino S, Howearth G. 1999. Thyroid health status of ammonium perchlorate workers: a cross-sectional occupational health study. J Occup Environ Med. 41(4):248-60

Lamm SH, Doemland M. 1999. Has perchlorate in drinking water increased the rate of congenital hypothyroidism? J Occup Environ Med. 41(5):409-11

Lang, L. 1995. Strange brew: assessing risk of chemical mixtures. Environmental Health Perspectives. 103(2). http://ehpnet1.niehs.nih.gov/docs/1995/103-2/focus.html.

Lawrence JE, Lamm SH, Pino S, Richman K, Braverman LE. 2000. The effect of short-term low-dose perchlorate on various aspects of thyroid function. Thyroid 10(8):659-63

Li FX, Squartsoff L, Lamm SH. 2001. Prevalence of thyroid diseases in Nevada counties with respect to perchlorate in drinking water. J Occup Environ Med. 43(7):630-4

Li Z, Li FX, Byrd D, Deyhle GM, Sesser DE, Skeels MR, Lamm SH. 2000. Neonatal thyroxine level and perchlorate in drinking water. J Occup Environ Med. 42(2):200-5

Mayes, Brian A., George L. Brown, Frank J. Mondello, Kevin W. Holtzclaw, Stephen B. Hamilton, and Arthur A. Ramsey. 2002. Dermal Absorption in Rhesus Monkeys of Polychlorinated Biphenyls from Soil Contaminated with Aroclor 1260. Reg. Tox. & Pharm. 35: 289-295.

Nessel C.S., Lewis S.C., Stauber K.K. and Adgate J.L. 1995. Subchronic to Chronic Exposure Extrapolation: Toxicologic Evidence for a Reduced Uncertainty Factor Hum. Ecol. Risk Assess. Vol. 1, No. 5, pp. 516-526.

Nosek, J.A., S.R. Craven, J.R. Sullivan, S.S. Hurley, and R.E. Peterson. 1992. Toxicity and reproductive effects on 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant hens. Journal of Toxicology and Environmental Health. 35: 187-198.

NRC. National Research Council. 1983. Risk Assessment in the Federal Government: Managing the Process. National Academy Press.

NTP. 1984. Toxicology and carcinogenesis studies of 1,3-butadiene (CAS No. 106-99-0) in B6C3F1 mice (inhalation studies). National Toxicology Program. NTP TR 288.

NTP. National Toxicology Program. 1989. Toxicity Study of Acetone, "Summary of Information on Acetone" (NTPTS2).

OSHA. 1995. U.S. Department of Labor. Occupational Safety and Health Administration. Code of Federal Regulations. 29 CFR 1910.1000, Table Z-l.

Park, RM. 2001. Perchlorate health effects in the epidemiological literature [with cover letter]. Cincinnati OH: National Institute for Occupational Safety and Health, Education and Information Division – Risk Evaluation Branch.

Piikivi, L. and U. Tolonen. 1989a. EEG findings in chlor-alkali workers subjected to low long term exposure to mercury vapor. Br. J. Ind. Med. 46: 370-375.

Piikivi, L. and H. Hanninen. 1989b. Subjective symptoms and psychological performance of chlorine-alkali workers. Scand. J. Work Environ. Health. 15: 69-74.

Piikivi, L. 1989c. Cardiovascular reflexes and low long-term exposure to mercury vapor. Int. Arch. Occup. Environ. Health. 61: 391-395.

Price, P.S., C.L. Curry, P.E. Goodrum, M.N. Gray, J.I. McCrodden, N.W. Harrington, H. Carlson-Lynch, and R.E. Keenan. 1996. Monte Carlo modeling of time-dependent exposures using a Microexposure Event approach. Risk Anal. 16(3):339-348.

Price, P.S., R.E. Keenan, and B. Schwab. 1999. Defining the interindividual (intraspecies) uncertainty factor. *Hum. Eco. Risk Assess.* 5(5):1023-1033.

Renwick, A.G. and N.R. Lazarus. 1998. Human variability and noncancer risk assessment – an analysis of the default uncertainty factor. Regul. Toxicol. Pharmacol. 27:3-20.

Restum, J.C., S.J. Bursian, J.P. Giesy, J.A. Render, W.G. Helferich, E.B. Ship and D.A. Verbrugge. 1998. Multigenerational Study of the Effects of Consumption of PCB-Contaminated Carp from Saginaw Bay, Lake Huron, on Mink. 1. Effects on Mink

Reproduction, Kit Growth and Survival, and Selected Biological Parameters. Journal of Toxicology and Environmental Health Part A 54:343-375.

Roy, T.A., J.J. Yang, A.J. Krueger, and C.R. Mackerer. 1990. Percutaneous absorption of neat 2,3,7,8-tetrachlorogdibenzo-p-dioxin (TCDD) and TCDD sorbed on soils. Toxicology 10(1):308.

SAB. 1997. An SAB Report: Guidelines for Cancer Risk Assessment. Sept. SAB-EHC-97-010.

SAB. 1998. An SAB report: Review of the health risk assessment of 1,3-butadiene. U.S. Environmental Protection Agency, Science Advisory Board, Environmental Health Committee. EPA-SAB-EHC-99-003. November.

Sample, B.E., D.M. Opresko, and G.W. Suter II. 1996. Toxicological Benchmarks for Wildlife: 1996 Revision. Risk Assessment Program, Health Sciences Research Division, Oak Ridge National Laboratory. ES/ER/TM-86/R3.

SCF. Scientific Committee on Food. 2002. Opinion of the Scientific Committee on Food on Fusarium toxins. Part 6: Group evaluation of T-2 toxin, HT-2 toxin, nivalenol and deoxynivalenol. European Commission Health and Consumer Protection Directorate-General. SCF/CS/CNTM/MYC/27 Final. February.

Schwartz J, Jewell NP. 2001. Gestational exposure to perchlorate is associated with measures of decreased thyroid function in a population of California neonates. Masters thesis submitted to UC Berkeley School of Public Health.

Sheppard, S.C. 1995. Parameter values to model the soil ingestion pathway. Environ. Monit. Assess. 34:27-44.

Simon, T.W. 1999. Two-dimensional Monte Carlo simulation and beyond: A comparison of several probabilistic risk assessment methods applied to a Superfund site. Hum. Ecol. Risk Assess. 5 (4): 823-843.

Stanbury JB, Wyngaarden JB. 1952. Effect of perchlorate on the human thyroid gland. Endocrinology 533-539.

Stehr-Green., P.A., E. Welty, and V.W. Burse. 1988. Human exposure to polychlorinated biphenyls at toxic waste sites: investigations in the United States.
Arch Environ Health. Nov-Dec;43(6):420-4.

Stanek, E.J., Calabrese, E.J., Barnes, R. and Pekow, P. 1997. Soil ingestion in adults--results of a second pilot study. Ecotoxicol Environ Saf. 36(3):249-57.

Stevens, K.R. L. and Mylecraine. 1994. Issues in chronic toxicology. In Principles and Methods of Toxicology, Third Edition. A. Wallace Hayes, editor. Raven Press, New York. Pages 673-695.

Suter, II, G.W. and C.L. Tsao. 1996. Toxicological Benchmarks for Screening Potential Contaminants of Concern for Effects on Aquatic Biota: 1996 Revision. Lockheed Martin Energy Systems. ES/ER/TM-96R2

Tucker, A.N., V.M. Sanders, D.W. Barnes, et al. 1982. Toxicology of trichloroethylene in the mouse. Toxicology. And Applied. Pharmacology. 62:351-357.

Uraih, L.C. and R.R. Maronpot. 1990. Normal histology of the nasal cavity and application of special techniques. Environmental Health Perspectives; Chemcials and Lung Toxicity, Upper Respiratory System. 85:187-209.

USACHPPM. U.S. Army Center for Health Promotion and Preventative Medicine. 2000. Standard Practice for Wildlife Toxicity Reference Values. Technical Guide No. 254. October.

Viscusi, W.K. and James T. Hamilton. April 1999. Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions (Working Paper 99-2). (Research support was provided under Cooperative agreement No. CR-823604-01 from the U.S. EPA Office of Policy, Planning, and Evaluation).

West, P.C., J.M. Fly, R.Marans, & F.Larhin. 1989a. Michigan Sport Anglers Fish Consumption Survey. Technical Report #1. (Natural Resources Sociology Research Lab)

West, P., J.M. Fly, R. Marans, and F. Larkin. 1989b. Michigan Sport Anglers Fish Consumption Survey. Supplement I – Non-Response Bias and Consumption Suppression Effect Adjustments. A Report to the Michigan Toxic Substance Control Commission. Natural Resource Sociology Research Lab Technical Report #2. September.

West, P.C., J.M. Fly, R.Marans, F.Larhin & D.Rosenblatt. 1993. 1991-1992 Michigan Sport Anglers Fish Consumption Study. Technical Report #6. (University of Michigan School of Natural Resources)

Wester, R.C., H.I.Maibach, L.Sedik & J.Melendres. 1993. Percutaneous absorption of PCBs from soil: In vivo rhesus monkey, in vitro skin, and binding to powdered human stratum corneum. J.Tox.& Env. Health 39:375-82.

WHO (1998). Environmental Health Criteria: Acetone. World Health Organization, International Programme On Chemical Safety. EHC 207. Geneva, Switzerland.

Wolf, O. 1976. Cancer Morbidity Amongst Chemical Workers From A Former Naphthalene Cleaning Plant [Translated From The German By The Uk Hse]. German Title: Kreberkrankungen Bei Chemiearbeitern Einer Ehemaligen Naphthalinreinigung. Dt. Gesundh.-Wesen.31:996-999. Hse Translation Number 15358 (A).

Wolf, O. 1978. Carcinoma Of The Larynx In Workers Engaged In The Purification Of Naphthalene [Translated From The German By The Uk Hse]. Zeitschrift Fur Die Gesamte Hygiene Und Ihre Grenzgebiete 24 (10: 737-739. Hse Translation Number 15329 (A).

Wu, C. and J. Schaum. 2000. Exposure assessment of trichloroethylene. Environmental Health Perspectives, Supplement 2 108:359-363.

Wyngaarden JB, Wright BM, Ways P. The effect of certain anions on the accumulation and retention of iodide by the thyroid gland. Endocrinology (1952) 50:537-549
Wyngaarden JB, Stanbury JB, Rapp B. The effects of iodide, perchlorate, thiocyanate and nitrate administration upon the iodide concentrating mechanism of the rat thyroid. Endocrinology (1953) 52:568-574

Xiao Li F, Byrd DM, Deyhle GM, Sesser DE, Skeels MR, Katkowsky SR, Lamm SH. Neonatal thyroid-stimulating hormone level and perchlorate in drinking water. Teratology 2000 Dec;62(6):429-31