14 June 2006
(via electronic mail)

Dr. Nancy Beck
Office of Information and Regulatory Affairs
Office of Management and Budget

Dear Dr. Beck:

In accordance with the Federal Register notice of January 17, 2006 (71FR2600), I herewith submit a public comment on the OMB Proposed Risk Assessment Bulletin. Per the requirements set out in the FR notice, my comments are in the attached Word file ('Rhomberg OMB Bulletin Comments.doc') and also follow below.

Thank you for the opportunity to comment.
Sincerely,

~ ~ Lorenz Rhomberg

COMMENTS ON THE OMB RISK ASSESSMENT BULLETIN

Lorenz Rhomberg, PhD

Principal
Gradient Corporation
20 University Road
Cambridge, MA
617-395-5552
lrhomberg@gradientcorp.com

In a January 17, 2006 *Federal Register* notice, the Office of Management and Budget's Office of Information and Regulatory Affairs put forth a *Proposed Risk Assessment Bulletin* for public comment. The proposed Bulletin sets goals and certain reporting standards that should be adhered to by Federal government agencies "to the extent appropriate" in "all agency risk assessments available to the public" (with some specified exceptions). It defines risk assessment as "a scientific and/or technical document that assembles and synthesizes scientific information to determine whether a potential hazard exists and/or the extent of possible risk to human health, safety or the environment." The goals and standards set forth in the proposed Bulletin are aimed at assuring that such assessments include forthright statements about scientific uncertainties in the projections or estimates of potential risks. In particular, there should be a frank and informative evaluation of plausible alternative assumptions and interpretations of the available scientific evidence, an objective assessment of their relative likelihoods of being true, and an assessment of their impacts on estimates of human health, safety, or environmental consequences.

Federal government risk assessment documents have large impacts not only on federal regulatory actions, but also on actions and approaches to public health and safety issues by state and local governments. Moreover, federal assessments affect public perceptions regarding the safety of chemicals and their uses, industrial activities, healthfulness of various foods and lifestyle habits, and more. Public and private choices must weigh the costs, risks, benefits, and fairness of alternative policies and courses of action. As a key and authoritative source of information about potential health and safety risks in numerous areas—chemicals; drugs; industrial activities and their products; consumer products; pesticides and food safety; alternative foodstuffs and lifestyle elements; transportation of people and goods; structural engineering of buildings, roads, and bridges; emissions from industrial production, power generation, and private automobiles; disposal of industrial and residential wastes, and more—it is important that the federal government should aim at providing clear and unbiased information so as best to inform not only its own regulatory and legislative choices, but also the choices and perceptions of society at large.

In the pages that follow, I present some comments on these questions and on the specific proposals of the *Proposed Risk Assessment Bulletin*. These comments are my own, but I acknowledge the support of the Halogenated Solvents Industry Alliance in underwriting my efforts to record and present them as a public comment to the Office of Management and Budget's Office of Information and Regulatory Affairs, in response to the solicitation of comments contained in the January 17, 2006, *Federal Register* notice.

I write as one who has had experience with federal risk assessments as a former federal employee producing such assessments (as a biostatistician doing chemical hazard and dose-response assessments for the US EPA, in its Office of Toxic Substances and Office of Research and Development), as a former academic teaching, doing research, and publishing on risk assessment methods (at the Harvard School of Public Health), and now as a consultant providing expert services on risk assessment matters to the private and public sectors (as a Principal at Gradient Corporation, an environmental science and product safety consulting firm headquartered in Cambridge, MA). Throughout my 20-year career in this field, I have grappled with the technical issues and sought ways better to express and communicate the scientific uncertainties inherent in assessing the potential for toxicological risks to humans from chemical exposure—both specifically in the context of the assessments for particular chemicals, and generally in my research and writing about science policies and chemical risk assessment methodology. I have been the subject of an expert elicitation project (*i.e.*, a source of information on scientific uncertainty in inference, a key technique in uncertainty analysis) and a developer of proposed statistical approaches for describing uncertainty in reference concentrations.

Accordingly, my viewpoint on the Proposed Bulletin focuses mostly on the risk assessments with which I am most familiar—those aimed at evaluating the potential of low-level chemical exposures to affect human health. I note that the applicability of the Proposed Bulletin is much wider, and I think many of my comments will have more general applicability. My experience is in trying to sort out conceptual issues and stumbling blocks, in finding practical ways to execute useful analyses of uncertainty in particular actual applications to chemicals, and seeking better ways to communicate such uncertainties to the technical audience for risk assessment as well as to the wider set of stakeholders. Such experience informs my comments.

GENERAL VIEWPOINT

In general, I support the objectives of the Proposed Bulletin. In particular, the goals of transparency and uncertainty characterization are important, and the standards for defining the scope and applicability of risk assessments are a valuable for setting the context of risk analyses. Many current risk analyses do a very incomplete job at characterizing the uncertainty in their projections. At the very least, each assessment should be required to assess its own assessment of uncertainty—that is, to examine what it was and was not able to do about characterizing uncertainty. This being said, there are a number of places in the Bulletin where clarification is needed. Different kinds of risk assessment (*e.g.*, those based on actuarial data and observation, those based on well-characterized causal systems such as engineered systems, and those based on testing of surrogate systems [as in animal toxicity testing] with major extrapolations resting on only partially understood underlying causal systems) will require rather different approaches to uncertainty characterization and can be expected to achieve clear uncertainty analyses to different degrees. How the Bulletin is to apply to different areas of risk assessment, and to assessments carried out for different purposes (general purposes, specific impact estimation, or screening, for example) is not well spelled out, and questions exist as to how compliance with the Bulletin's standards is to be judged. As a mandate rather than guidance, the Bulletin does not give much help in suggesting how its requirements are to be addressed, and it is unclear what analyses will be considered sufficient. There are many available techniques for quantitatively characterizing uncertainty, but many have yet to win wide acceptance and they have not been standardized, so one can anticipate challenges in implementing the Bulletin in a way that meets general acceptance. In my more specific comments, I address some of these methods and the issues that may arise. I find that uncertainty analysis is indeed possible, certainly valuable, but not necessarily as easy or unambiguously defined as the Bulletin suggests. I would recommend a revision of the Proposed Bulletin that is more specific about different applications, both in terms of settings for risk assessment, the main sources of uncertainty, and the methods that appear

appropriate to characterize them. One specific challenge will be that the Bulletin places appropriate focus on articulating the specific risk management application of an assessment as a context within which to judge the uncertainties and their importance to the ultimate decision—but many of the chemical hazard assessments to which the Bulletin is to apply are hazard characterizations and dose-response analyses that are intended to be generally applicable to a variety of risk management applications.

SUMMARY OF MAIN POINTS

I wish first to summarize my main points. I will elaborate on these further on.

- Many current risk assessments do at best a rudimentary job at characterizing uncertainty, but the uncertainty may be considerable.
- Current health risk assessment methods for chemical exposures focus on upper bound risks, and such analyses poorly serve many current risk management problems. Such assessments may distort comparative risk analyses, substitution decisions, and may result in poorly set priorities that focus on lesser understood risks rather than those of most impact on public health. Benefits assessments, cost-effectiveness analyses, and cost-benefit analyses can also be distorted by considering only upper-bound risks of unknown and differing degrees of conservatism.
- A fuller and more forthright characterization of uncertainty would benefit risk assessment and result in sounder risk management decisions. It is important to move ahead the treatment of uncertainty in federal risk assessments.
- The Bulletin is a mandate for more extensive reporting of uncertainty, but it is not guidance for how to conduct the analysis, nor specifically how to do the reporting. Implementation of the Bulletin is therefore critical to determining whether it will achieve its aims, and challenges to successful implementation must be acknowledged.
- The Bulletin's proposals for goals related to problem formulation, scope, and the prior statement of information needs and objectives are sound and will help to provide the proper context for determining methods and interpreting results. Relating the analysis to the questions raised by the risk management applications being supported helps ensure that methods and the presentation of results are appropriate to their eventual use.
- That being said, the Bulletin proposes to apply its standards to those parts of risk assessment—in particular, to hazard characterization and dose-response assessment—that are often undertaken separately from any specific risk-management application (e.g., in EPA's IRIS database) and indeed are intended to be generally useful to a variety of applications. The Bulletin should clarify how the standards for problem formulation and the mandated connection to the application are to be carried out in such cases.
- The goals of transparency and peer review are important and laudable.
- The standard for identifying and examining the consequences of critical assumptions is important and laudable. The Bulletin acknowledges that some assumptions may have to be analyzed qualitatively and with sensitivity analysis, rather than with a full quantitative description of the uncertainties involved.
- That being said, the Bulletin mandates quantitative analysis of uncertainties to the extend possible and appropriate. A basis for judging what is possible and what is appropriate in different cases is not clear, and the Bulletin as currently phrased raises questions about what will be considered to be in compliance. Because of the proviso for certification, this will be a critical question that should be clarified.

- The scope of an uncertainty characterization, in terms of thoroughness, depth, and completeness of characterization of all identifiable sources of uncertainty, will be a critical question. The Bulletin calls for "appropriate" efforts and depth of analysis, but the definition of what will suffice is unclear. The Bulletin needs to be more specific about how one is to judge the types and levels of uncertainty analysis that should be treated as required for different types of risk assessment. Some language in the Supplementary Information seems internally contradictory regarding what analyses apply to screening assessments.

- The readiness with which uncertainties can be expressed in the form of quantitative distributions varies considerably among types of risk assessment. Failure analyses of engineering systems are readily expressed in this way, since the system and its failure modes are well understood and expressible in terms of probabilities of readily definable events, which probabilities can be estimated from experience and experiment. Actuarial assessments of risk are similarly able to create such distributions, because the question is essentially a statistical one based on frequencies of directly observed events. Such data are amenable to meta-analyses. Projections of human low-dose health risks from high-dose animal studies, however, entails uncertainties that are more about the extent of validity of models of inherently difficult to observe and poorly understood underlying biological processes, about the validity of extrapolations and analogies between animal and human biology—in short, about the validity of theories of unobserved or unobservable phenomena. Such uncertainties are challenging to express in terms of distributions.

- That being said, there are methods that have been applied to such problems. I myself have been a participant (as an elicited expert) in an expert-judgment analysis of carcinogenic potency for chloroform (Evans et al., 1994. Regul Toxicol Pharmacol 20(1 Pt 1):15.) and I have published a distributional approach to the Reference Concentration for ethylene oxide (Evans *et al.*, 2001. Risk Anal 21(4):697).

- The Bulletin should be construed so as to encourage such analyses. To mandate them in the short term, however, may be problematic, because such analyses are resource intensive, do not have universally agreed-upon methods, and may be more difficult for cases with less developed databases than those available for the extant examples. There should at least be a mandate that each risk assessment should assess the state of its own uncertainty analysis. (This is what the Certification appears to require, but that could be clarified.) That is, an assessment needs to identify the main uncertainties and evaluate how well it has dealt with characterizing them.

- Consideration of "negative" studies, as the Bulletin proposes to mandate, would improve risk assessment, but in many cases there is no clear way to do so quantitatively. For example, lack of concordance in specific effects among animal studies, and between animal studies and human studies, could be considered as negative results in some studies for particular endpoints.

- The quantitatively most important uncertainties in chemical hazard assessment are the ones least readily expressible in quantitative form—the weight of evidence for causation of particular endpoints, the applicability of high-dose outcomes to low exposure levels, and the implications of animal experiments for human risk. The Bulletin should take care not to focus attention on relatively minor sources of uncertainty simply because they are more quantifiable. This could lead to conventionalized uncertainty analyses, in which by precedent and emerging practice some fairly standard analyses are carried out and represented as "the" uncertainty in a risk assessment, while missing the main issues and presenting a misleading picture of the actual degree of uncertainty. The Bulletin could require listing sources of uncertainty in order of how large an impact they are understood to have, or it could require separate listings of major and minor uncertainties.

- The Bulletin presumes that quantitative characterization of uncertainty can readily be expressed in terms of a single "central estimate" and distinct upper and lower bounds—in the manner of a familiar mean and confidence interval—and such a simple image may engender false expectations about what can be done and should be done in a quantitative characterization of uncertainty. First, the sources of uncertainty are multiple, and to meet the goals of transparency and communication about uncertainties, it is necessary to characterize the component contributors as well as any overall uncertainty. The interpretation of different factors—especially the distinction between uncertainty and inter-individual variability—are important in appropriate interpretation of uncertainty analysis and in proper application of its results to risk management problems. "Central" estimates are often not near the center of their distributions, because the distributions are skewed. The Bulletin should acknowledge these statistical factors.

- The Bulletin (in its Supplementary Information) uses the terms "central" and "expected" synonymously. This would appear to mandate expected values as the desired estimate. It should be clarified that expected values are means of distributions and that they need not be—indeed they are often not—the most likely values nor ones near the middle of a distribution. (As an example, model averaging, as mentioned in the Bulletin, would, if faced with one low-dose extrapolation, with a 90% probability of being true and predicting zero risk, and an alternative model deemed to have a 10% chance of being true and predicting 10^{-4} risk, calculate an expected value of 10^{-5} risk.) The expected values need not even be possible values for distributions of discrete states. Other measures of central tendency can also be useful in communicating about uncertainty and its impact on a risk assessment's findings, and the Bulletin should not be unduly restrictive. The Bulletin should sharpen its discussion and definition of what it intends to mean by "central" estimates.

- In any uncertainty characterization, there are choices for the analyst to make regarding which contributory factors are included and which excluded, which ones are to be treated quantitatively and which qualitatively, which among several bases are to be used to characterize particular factors (especially when the degree and nature of the uncertainty is based on models, analogies, or projections rather than observation of many cases), whether the analysis includes recognized factors or an allowance for the possible role of heretofore unrecognized factors, and so on. These choices are made based on available information, codification of the understanding (or lack of understanding, or alternative theories) about underlying pharmacokinetic and pharmacodynamic processes, and the needs of the risk management process being supported. It may be dangerous to specify that one characterization of "the" uncertainty is possible or even desirable. To do so risks conventionalization of the analysis (to define what is expected and to make analyses "comparable" across assessments) in a way that may distort the main purposes and goals of the Bulletin. In order to address this point, the Bulletin should focus more on characterizing specific components of uncertainty, and not imply that the only outcome of importance is the "bottom-line" characterization of overall uncertainty from all sources. It should take care to give legitimacy to qualitative as well as quantitative analyses of factors contributing to uncertainty, instead of implying that everything should be quantified and combined into a single uncertainty interval if at all possible. Such an emphasis on the bottom-line may undermine the goals, purposes, and utility of the Bulletin's mandates.

- In view of the previous point, defining the scope and specifics of each assessment's uncertainty analysis is a critical aspect, and the Bulletin provides little basis to understand how the appropriateness of the scope is to be judged case by case.

- The Bulletin mandates that health and safety analysis endpoints be judged to be adverse, and that an assessment should evaluate such adversity. While it is important to attend to the adversity of endpoints in interpretation, and to interpret public health protection in terms of avoidance of such

adverse effects, it is also true that analysis of precursor endpoints and markers of underlying processes can greatly enhance a risk assessment, its interpretability, and the characterization of its uncertainty. The Bulletin should clarify that the analysis of markers is not being discouraged.

- The Bulletin says little about how uncertainty analysis results are to be used in risk management. This may be beyond the scope of the current Bulletin, but it will be critical to the successful implementation of the Bulletin's mandates. Risk assessment needs to be responsive to the questions being asked of it, and if the purpose is to communicate uncertainties to risk managers in a way that improves their decision-making, the interactions between assessment and management are important.
- In sum, it is important to move ahead on characterizing uncertainty in toxicity assessment. Doing a fuller job of acknowledging and evaluating sources of uncertainty can draw on a number of existing tools. A good analysis is possible, but not necessarily easy. It is hard to be comprehensive and universal about how to characterize uncertainty, and there are no self-evident "standards" for analysis that can be applied in all cases. Different implementations will be different. This raises questions about what should be understood as complying with the Bulletin's mandates, and a fuller explanation of what should be expected under what circumstances would improve the likelihood of successful implementation.

DISCUSSION

Many current risk assessments regarding potential health effects from chemical exposures lack a sufficiently complete and forthright characterization of uncertainty. That is, there is indeed a problem to address. Some guidelines for chemical health risk assessment provide general acknowledgement of uncertainties inherent in the risk assessment process and they may list the major assumptions, but these are rarely applied in any specific way to the assessment of a particular chemical. Moreover, the impact of alternative assumptions is rarely evaluated. To varying degrees, health effects assessment documents may consider alternative datasets, dose-response models, cross-species extrapolation methods, and theories of modes of action, but these considerations are buried in the body of the text, and in the end a single result is chosen, which becomes all that most users of the risk assessment ever see or consider in risk management analysis.

Upper Bounds

Most current risk assessments regarding potential health effects from chemical exposures are focused almost solely on defining "upper bounds" on risk (or "lower bounds" on an acceptably safe dose). Most identified uncertainties in the inferential process are treated by considering the single alternative thought to be unlikely to underestimate risk. (For brevity, I will refer to the "upper bound" nature of risk estimates, but in doing so I intend also to refer to lower bounds on acceptable doses.) At least historically, there were legitimate reasons for this approach, but such assessments do not serve many modern risk management purposes well (as discussed below). In guidance and policy documents, and even in the body of assessment documents themselves, the upper bound nature of the findings is often fully acknowledged, but in practice they are used as though they were fully ascertained properties of the chemicals in question.

Despite the upper-bound nature of current chemical quantitative risk assessments, users (i.e., risk managers) tend to use the values as though they are true, (which may hamper sound decision-making). In part, this is ascribable to there being no specific alternative risk values offered for use. There is currently

a "chicken and egg" problem—risk managers cannot explore the impacts of uncertainties and central estimates on their risk-management decisions because they have no quantitative basis to do so in the form of quantitatively characterized uncertainty in risk; but lacking such information, they follow risk management analyses that do not require such uncertainty characterization. If the risk management process demanded more characterization about the uncertainty distributions of the risk estimates they are being provided, the risk assessment process will more readily rise to the occasion.

Reasons for Reluctance to Fully Characterize Uncertainty

Lack of full characterization of uncertainty in current assessments stems from several sources: (a) historical roots of risk analysis and the weight of precedent; (b) the mandates of most environmental laws; (c) fear of undermining the basis for appropriate regulatory action; and (d) inherent difficulties in fully characterizing uncertainty. Regarding (a), current risk assessment methods are elaborations of the first methods employed in early risk assessments, which were aimed at evaluating whether existing very low exposures (e.g., food additives) were comfortably below levels of concern for human health impact even in the face of uncertainties in the inferences. There has been pronounced "mission creep" in risk assessment from this early focus on evaluating the safety of already low exposures as they are found. The next step is to ask how much higher exposures could be and still be considered assuredly safe (e.g., in setting limits on permissible exposure to pesticides). This puts pressure on the scientific justifiability and confidence in the magnitude of the uncertainty factors—instead of approximate margins of safety, the idea became institutionalized that the chosen margin divides exposures that are comfortably safe from those that cannot be assured of being safe. The next change was to apply the assessment to exposures that exceed levels deemed comfortably safe, providing evidence of the need for mitigation of such exposures (e.g., waste-site evaluations, groundwater cleanups). Most recently, risk assessment is called upon to characterize the actual likelihood of ill effects from higher exposures, and in a way that can be used in cost-effectiveness and cost-benefit analysis and in regulatory impact analysis. Despite these changes, the fundamental structure of chemical health hazard assessment has remained the same (although there have been elaborations and refinements), and it is not surprising that it has had difficulty fully meeting new challenges. The fundamental approaches of safety factors and conservative extrapolations, formulated for the first risk assessment applications, have become institutionalized.

Regarding (b), many environmental laws contain little in the way of specific guidance for risk assessment. Instead, they contain generally phrased mandates such as to "ensure protection of the public with an adequate margin of safety" or to ensure "a reasonable probability of no harm." Early risk analysis was designed to address these concerns for bounding estimates of acceptable levels of exposure. For these laws, such bounding estimates are still required, but a fuller characterization of uncertainty could still improve the determination of such bounds and would provide important insight into the protectiveness of resulting regulations and what the "adequate margin of safety" really constitutes in practice. Nonetheless, the perception that the laws do not allow probabilistic approaches has hampered their development.

Regarding (c), regulations cannot set standards that are "arbitrary and capricious," and some have feared that if risk findings are not expressed as definitive numbers, but only as distributions of possibilities, then any regulatory standard based on them might be ruled arbitrary. It is not clear, however, that such decisions are any less arbitrary than current practice in which upper bounds of unknown conservatism are applied, possible carcinogens are treated exactly as known human carcinogens are treated, and a one-in-a-million risk, which could never be demonstrated but arises only from accepting orders-of-magnitude extrapolations of uncertain validity. It would seem that risk estimates of specified and specifically estimated conservatism would be less arbitrary than current practice, not more. Another factor is that the regulatory apparatus is intentionally designed to make the decision-making process rely

on objective analysis findings of fact, and consistent rules—that is, it is intentionally institutionalized—and specifically to avoid giving authority to individuals to make decisions based on their own opinions or interpretations. It is a way of limiting individual power and discouraging politicization. Environmental laws are written to mandate regulations based on findings of fact on the somewhat disingenuous presumption that any competent scientist would come to the same conclusions. In fact, regulatory actions can be adversarial, and parties that are in fact arguing about values and interests recast their positions as arguments about science. To the degree that the regulatory bureaucracy admits uncertainty in the science, there is a perception that it may tend to undermine the authority to take what may be needed action, providing a de-motivation to undertake fully elaborated uncertainty analysis.

Regarding (d), it must be admitted that a full characterization of uncertainty is a difficult undertaking. Some issues will be discussed further below, but it is clear that there is no obvious and straightforward universal method that would create a full quantitative characterization of uncertainty that would be acceptable to all parties. The sources of uncertainty are multiple, understood or even recognized to varying degrees, susceptible in varying degrees to quantitative description, and of different levels of concern to different stakeholders. Once one announces that one has tried to characterize "the" uncertainty, it opens numerous avenues of criticism, alternative interpretations, and arguments about alternative means of analysis. Thus another factor tending to discourage uncertainty analysis at federal agencies is the effort and resources needed to do it properly, and the further effort needed to defend it in the face of critical commentary to which there is no definitive solution.

Assumptions as Simplifications

Many of the assumptions of risk assessment are simplifying assumptions and the alternatives do not always have discrete states that can be associated with specific identifiable consequences. For example, the assumption that all exposures can be compared on the basis of a daily average intake (even if some are long exposures to low levels and others are brief exposures to high levels) is a simplification that permits a single dose metric to apply to a variety of situations (including those from which one is extrapolating and those to which one is extrapolating), but it is usually uncertain whether this is really true. The alternative, that exposures cannot be time-averaged, is non-specific—it can happen in an infinite number of alternative ways—and there may be little information with which to characterize the consequences of assuming that the simplification does not apply.

Benefits of Quantitative Characterization of Uncertainty

There is considerable benefits of expressing uncertainties as quantitatively as possible. Current "upper bound" methods play down the considerable uncertainties and, by being presented as definitive numbers, are often inappropriately interpreted as estimates of actual (rather than possible) impacts. Current methods are elaborations of an older "safety assessment" approach, which aimed at documenting the lack of concern for very low exposures but was not designed to address actual impacts of higher exposures. This older approach is ill suited to many modern risk management problems, where one seeks to balance costs and benefits, to compare potential risks of alternative chemicals or alternative control measures, or to evaluate the likely public health benefit of proposed chemical-control regulations.

Even when one aims at regulation that is conservative in the face of uncertainty, it is important to know how conservative one is being and whether extreme control measures buy meaningful increases in the assurance of protection compared to more moderate measures. When one is doing comparative risk assessment, comparisons among chemicals with different degrees of conservatism can be biased, and ignoring uncertainty can obscure analysis of whether trade-offs, substitutions, or alternative technologies yield dependable improvement.

Thus, better characterization of uncertainty in risk assessments is important, and doing so is not in itself aligned with any particular risk management approach or regulatory agenda. In particular, better characterization of uncertainty would be beneficial to better execution of the decision-making mandates of a variety of environmental and occupational protection laws, not just those entailing cost-benefit balancing. The real aim is transparency and forthrightness about the limits of our knowledge, not any particular way of dealing with those limits.

Methods for Quantitative Uncertainty Assessment are Available

A variety of methods have been developed to estimate and express the degree of uncertainty in quantitative risk assessment of chemical hazards, including meta-analysis, sensitivity analysis, expert-judgment methods, Monte Carlo analysis, and others. The Bulletin has a rather brief discussion of these, and would benefit from a fuller explanation of the kinds of methods it foresees as fulfilling its mandates. Some of the references to examples and key literature are a bit dated, and newer examples, crafted in the context of current concerns, would be valuable to cite. Some discussion of the different contexts in which various approaches would be most appropriate would also be helpful, along with forthright discussions of the limitations of the available methods and an examination of pitfalls or caveats that ought to be borne in mind. The aim should not be to write a tutorial or give specific technical guidance, but to acknowledge that some clarity ought to be given about what kinds and extent of uncertainty analysis is being mandated by the Bulletin.

Challenges Exist to Real Fulfillment of the Bulletin's Goals and Mandates

The benefits of better uncertainty analysis in risk assessments is clear. Nonetheless, it is important to acknowledge the challenges, difficulties, and limitations involved. Failing to do so risks unintended consequences. Poor analyses or poor interpretations of analyses could become bureaucratically ensconced as standard approaches for fulfilling the Bulletin's mandates. Conventional interpretations and approaches could make it difficult to add information or modify assessments of uncertainty in the future. Exactly this kind of hardened conventional methodology and interpretation has happened for the "upper bound" of cancer potency analysis and for the conventional uncertainty factors in non-cancer risk analysis, and these conventions, supported by precedent more than by information, are hard to change with case-specific science. Also, failure to acknowledge the limitations of uncertainty analysis could lead to bad decisions, could work against proper interpretation of case-specific data, and could lead to cynicism about the whole enterprise of examining uncertainty in risks during the decision-making process.

The Bulletin seems somewhat naïve about the challenges, as though central estimates, well-defined bounds, and probability distributions are all readily available to an analyst, with the only need being to mandate their presentation in the final risk analysis document. The implicit mental model of statistical confidence limits for fitted parameters is only a small part of the overall uncertainty problem, and the really influential uncertainties are not nearly as well defined or easily described. The most challenging uncertainties are the more qualitative ones, such as whether a compound is a carcinogen or not, which among many alternative datasets should be used for dose-response analysis, what uncertainties are involved in the major extrapolations (animal-to-human; regular experimental exposure regimes to intermittent or fluctuating real-world exposures, *etc.*), how reliable is a theory of mode-of-toxic-action vis-à-vis conceivable alternatives, and so on. Expert judgment can be used in such matters, but it raises questions about how to choose the experts, whether government agencies can delegate decisions to expert panels, how to ensure that multiple stakeholder groups are satisfied with the legitimacy of the analysis, and the considerable expense involved. I don't think the potential difficulties justify ignoring or

downplaying the uncertainty issue—indeed, it is to be transparent about such matters that is the chief need—but it should not be assumed to be a simple, entirely objective analytical process with unambiguous or unchallengeable answers that is available off-the-shelf.

The Importance of the Statement of Scope

A key aspect is the definition of the *scope* of the uncertainty analysis, because what is possible to address (and with what methods) will vary from case to case. For example, in characterizing inter-individual variability one may be relying on general empirical observations of the extent of variability in response to other chemicals or one might have chemical-specific information on the *causes* of potential variation in sensitivity among people, and in yet others there might be speculation about the existence of ultrasensitive people based on sparse, inconclusive data that are nonetheless sufficient to suggest possibilities not seen for chemicals in general. In some cases, the uncertainty would be about the "normal" degree of variation in responsiveness, while in others, specific causes of idiosyncratic reactions might be known. The question about what fraction of the whole US population would be protected by an exposure limit, and whether there are actual or only hypothetical sensitive individuals may not admit to consistent answers across chemicals.

Negative Studies

The Bulletin mandates consideration of "negative studies." While this is very important, it is also laden with questions about which negative studies are to be considered and what their apparent negativity means for the overall risk analysis. In a carcinogenicity assessment, for instance, one might want to consider the fact that hamsters show no tumors in lifetime exposure while mice show liver tumors and rats show kidney tumors. But should one also consider the negative response of rats for liver tumors and mice for kidney tumors? How about the negative response of all species for thyroid tumors? If the existence of apparent species-differences in target-organ response is used to affect the estimate of uncertainty in the application of animal data to humans, are the various laboratory animal responses to be seen as a sample of all possible mammalian responses (with humans being just another mammal) or are specific relevancies of particular animal tumors to humans (as judged by other chemical assessments) to be estimated, or are the apparent differences to be analyzed as a product of different dose levels (interacting with dose-response analysis) or as consequences of different statistical power across studies, or of mortality-pattern differences, or of animal-feed differences? Again, the point is not that complication should make us avoid asking these questions, but rather that there is no single kind of "uncertainty analysis" that one can mandate that would answer such questions in an unambiguous, universally acceptable and operational way across all compounds, and thus it is hard to know what the Bulletin actually mandates and whether specific analysis would or would not be in compliance.

Questions About What Will Constitute Compliance

The Bulletin makes much of distinguishing "influential" assessments (with added requirements) from others, notes that a full uncertainty characterization might not be needed in some "screening" assessments, and exempts certain assessments supporting licensing or site-specific actions. The definitions of different types are not very clear and will generate questions about applicability of the Bulletin's requirements. The Bulletin stresses that "appropriate" levels of uncertainty analysis need to be done and makes some concession to practicability, but again with little basis for knowing what distinctions are to be made when. These will need to be clarified.

The Bulletin notes that sensitivity analysis to alternative assumptions or choices of data needs to be conducted. This is a very worthwhile practice and does a lot to spell out qualitative uncertainties.

There are pitfalls, however, and they are well illustrated by the 2001 draft EPA reassessment of trichloroethylene (TCE). That assessment considered a considerable number of alternative datasets and models as candidates for the basis of human cancer risk potency estimation. This examination of alternatives could be considered an uncertainty analysis and an exercise in the kind of sensitivity analysis the Bulletin envisages. But the EPA TCE reassessment does not go on to use the perspective generated to support any decisions about how human cancer potency is to be treated, simply presenting the range of alternatives and advising the user to use the "most appropriate" one for a particular application, without any discussion or guidance as to how to do so. In fact, all the alternatives are not equally plausible or scientifically supported, and it has been left to others (including a paper by Lewandowski and Rhomberg, 2005. Regul Toxicol Pharmacol 41(1):39) to try to articulate a basis for the choice. Moreover, the alternatives are all alternative upper bounds, and despite the 20-fold range do not give a real picture of the true uncertainty in projections of low-exposure cancer risk from human TCE exposures. Many relevant factors are left out, and there are no "central" estimates. The practical consequence is that various EPA regulatory program offices, as well as various state agencies, have had to grapple with completing the consequences of the "uncertainty analysis" for themselves, often without the resources and expertise necessary. There has been a predictable reliance on the upper end (*i.e.*, the most conservative) member of the range as an alternative that cannot be gainsaid as insufficiently protective, resulting in an unwarranted increase in the conservatism of the overall assessment. Affected private parties and the citizenry at large have been left with no basis to predict how a government assessment affecting their interests might end up being conducted. This is not the kind of uncertainty analysis that is helpful. Somehow, the expression of uncertainty in the assessment needs to be coupled with the decision-making process so that the assessment of uncertainty illuminates decision-making, rather than paralyze it (as has happened with TCE).

Adversity of Endpoints and the Value of Markers of Underlying Processes

The Bulletin calls for assessments to identify which response endpoints are "adverse" (*i.e.*, indicative of frank harm) and which ones are merely intermediate endpoints that show the operation of some presumably relevant process that may later on (or at a higher level of activity) lead to adverse effects. In particular, mere markers of accommodative processes or the induction of defenses should not be used as an effect in itself. While this is important, it is also important to note that the adversity question itself can be an uncertain call that needs assessment. Moreover, such observations of precursor effects or markers underlie much of the basis for estimating the degree of uncertainty, of inter-individual sensitivity, of the reliability of a theory of mode of action, and other highly relevant aspects of an assessment. The Bulletin should clarify that the assignment of adversity to endpoints is not intended to preclude these uses of precursor data. In addition, adversity comes in degrees, and projecting adversity in humans from responses seen in animals itself entails a good deal of uncertainty that will need to be analyzed.

Summary

In sum, the Bulletin calls for a greater degree of transparency in dealing with the uncertainties inherent in risk assessment in analyses carried out by Federal agencies. Such transparency and uncertainty analysis are very beneficial. In mandating such analyses, however, a lot of questions are raised about what specific analyses are to be required for which specific kinds of assessment. What kinds and levels of analysis will be considered sufficient to fulfill the Bulletin's mandates is not very clear. The specifics of the activities mandated by the Bulletin are not very explicit, and the tone and substance of the guidance fails to acknowledge the difficulties and pitfalls that will arise in practice.

In the end, risk assessment is a tool for risk management decision-making. What analyses are appropriate and useful will be dictated by the decisions to be made. Risk managers can bear in mind the

limits and incompleteness of uncertainty analyses that accompany risk assessments so long as these are transparently and fairly portrayed. The difficulty of doing this by mandates on the content of risk assessments is that this connection of the risk analysis to the decision-making is broken. The kind of general-purpose, comprehensive, standards-meeting evaluation of uncertainties that the Bulletin envisages may be hard to accomplish in practice, and what is accomplishable may not always serve the risk management process well. Better and more transparent evaluation of uncertainties in chemical risk assessments are important goals—I support them and have worked to develop means to meet them—but a more effective means of bringing this about might be to encourage the risk management process to demand more uncertainty evaluation from risk assessment than to mandate that all risk assessments conduct some vaguely specified uncertainty analysis.