John G. Haggard, Manager
Hudson River Program

General Electric Company
320 Great Oaks Office Park, Ste: 323
Albany, NY 12203
Fax: (518) 862-2731
Telephone: (518) 862-2739
Dial Comm: 8* 232-2739
E-Mail:John.Haggard@corporate.ge.com
Pager: 518-484-3177

April 28, 2000

Sean Kennedy, Ph.D.
Ecotoxicology Consultant
80 Saddlehorn Crescent
Kanata, Ontario K2M 2B1 Canada

Re: Hudson River Ecological Risk Assessment

Dear Dr. Kennedy:

I am writing regarding the upcoming peer review of the Hudson River Baseline Ecological Risk Assessment (BERA) in which you are to participate. General Electric Company (GE) believes there are several fundamental problems with the BERA, and attached to this letter are two short papers summarizing our view. I hope that you consider these papers and find them useful in your deliberations.

The most fundamental problem with the BERA is its lack of transparency and clarity. On March 21, 1995, EPA Administrator Carol Browner issued guidance regarding risk communication.[1] In this guidance, Administrator Browner described the need for objective and balanced communication to the public in all risk assessment documents and policy decisions. She stressed the need for:

- Transparency and clarity in risk communication and management, and
- Consistency and reasonableness in the implementation of risk assessments.

The call for transparency and clarity in the decision-making process refers to the need for full and honest communication of all objectives, assumptions, uncertainties, and criteria that go into the development of an EPA policy decision. On this topic, the Administrator wrote:

> *This means that we 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 as they are made throughout the risk assessment and risk management processes.*

[1] This memorandum is available for viewing or download on the EPA web page at http://www.epa.gov/ordntrnt/ORD/spc/rccover.htm

The BERA falls short of both of the Administrator's clear and unambiguous directives. The risk characterization process and presentation is neither clear nor transparent.

- The BERA contains a detailed discussion of the mechanical steps involved in the evaluation of the various measurement endpoints, but fails to discuss the underlying principles, assumptions, and thought processes implicit in the interpretation of the findings. The BERA purports to use a "weight of evidence" approach in the characterization of risk. EPA did not actually apply this approach, however, instead relying almost exclusively on "toxicity quotients" (TQs). The TQs used by EPA, moreover, are derived from measured or predicted concentrations of PCBs and literature-derived toxicity reference values (TRVs). Virtually no information is presented documenting the actual condition of the ecological resources of the Hudson River Valley or demonstrating any actual harm caused by PCBs. In essence, EPA has redefined the term "at risk" to mean "not proven to be risk-free" using conservative models and assumptions.

- The BERA also mischaracterizes and greatly understates the uncertainties associated with the assessment. The "Uncertainty Analysis" presented as Chapter 6 is largely made up of a discussion of data quality and a generic catalog of the types of uncertainties that may exist in risk assessments in general. There is no attempt to summarize or even roughly quantify the net uncertainty associated with the assessment endpoint evaluations or risk assessment conclusions. With the exception of some food web model input parameters, no uncertainties associated with the various endpoints are quantified. The reader is left with a false impression that uncertainty is relatively low, when in fact, it is orders of magnitude for many assessment endpoints. In general, the assessment contains a large number of assumptions, uncertainties, and implied policy decisions that go unstated and unexplained.

- Policy decisions contained within the BERA are disguised as scientific decisions, in direct conflict with the Administrator's memorandum. The "uncertainty factors" used in the derivation of the TRVs are actually safety factors that increase the conservatism of the assessment without directly addressing scientific uncertainty.

The result is an assessment that is unreasonable and results that are unrealistic. EPA did not rigorously evaluate all available data to reach reasonable and clearly justified conclusions. Data not considered by the Agency include recent studies demonstrating that one of the endangered species determined by EPA to be "at risk," the shortnose sturgeon, has actually greatly increased in abundance throughout the lower Hudson. Another endangered species that was determined to be "at risk," the bald eagle, is now successfully nesting and reproducing in the Hudson River Valley for the first time in decades.

GE believes that the problems in the BERA make the document of little use for the risk manager. The attached papers provide more detail about these problems. I hope that you review them before reaching a final conclusion about the BERA. Please do not hesitate to contact me if you have any questions.

Yours truly,

John Haggard/bg

John G. Haggard

JGH/bg

Attachment

cc: Alison Hess, U.S. EPA
William McCabe, U.S. EPA
Douglas Tomchuk, U.S. EPA
Douglas Fischer, U.S. EPA (ORC)
Steven Sanford, NYDEC
Ron Sloan, NYDEC
Bob Montione, NYDOH
Tom Brosman, NOAA
Jay Field, NOAA
Sharon Shutler, NOAA
Lisa Rosman, NOAA (New York)

Assessment and Measurement Endpoints/Risk Characterization/Uncertainty Analysis

The Baseline Ecological Risk Assessment (BERA) peer review panel was asked to consider whether the combination of measurement endpoints selected by the EPA supports the weight of evidence (WOE) approach used in the BERA (Specific Question 3) and whether the risk characterization adequately characterizes risks to ecological receptors (Specific Question 8). These issues are closely related. EPA's failure to select proper measurement endpoints and to evaluate the various lines of evidence and integrate information in a transparent, objective framework leads to an inadequate characterization of risks to ecological receptors in the following ways:

- The risk characterization does not adequately characterize risks to ecological receptors because the suite of measurement endpoints used in the BERA has little relevance to population effects and as such, is insufficient.

- The risk characterization should have relied on a WOE approach that evaluated endpoints and assigned differential weights to them based on the ecological relevance of each endpoint, the amount of site-specific data used to characterize each endpoint, and the relative uncertainty.

- Given the complexity of the site and the abundance of data, the BERA should have relied heavily on empirical site-specific data on the status of local populations and evidence for adverse effects of PCBs on those populations.

- EPA cites no site-specific empirical data that indicate the presence of adverse effects of PCBs on the survival, growth, or reproduction of local populations. Moreover, GE is not aware of any empirical data that document these types of cause and effect relationships in Hudson River biota.

The risk characterization does not adequately characterize risks to ecological receptors because the suite of measurement endpoints used in the baseline ecological risk assessment (BERA) is insufficient.

The BERA is based principally on "toxicity quotients" (TQs), i.e., comparisons between measured or modeled exposure concentrations and concentrations believed to be potentially harmful to organisms. Conservative, "screening-level" data and assumptions are used to define both the exposures and the effects. Screening-level data and models, such as those used by EPA, are deliberately designed to be conservative, i.e., to minimize the possibility that any potential adverse effects will be missed. As such, they overstate the actual effects of most chemicals at most sites. With the exception of the benthic invertebrate community survey and the tree swallow survey, all of the measurement endpoints used by EPA in the Hudson River BERA are measures of exposure or generic toxicity benchmarks: sediment-quality guidelines, water-quality criteria, and toxicity reference values (TRVs) derived from conservative single-species toxicity tests. Water-quality criteria and sediment quality guidelines should be used only to infer potential for risk, as in a screening-level assessment (EPA 1997). The *Ecological Risk Assessment Guidance for Superfund* (EPA 1997) explicitly states that decisions to require remedial action based solely on screening-level data "would not be technically defensible." A scientifically defensible ecological risk assessment should be based on the methods described below, not on TQs.

A wide variety of techniques for measuring and characterizing ecological risks at contaminated sites are described in EPA's *Guidelines for Ecological Risk Assessment* (EPA 1998) and *Ecological Risk Assessment Guidance for Superfund* (EPA 1997). These include:

- Measuring the abundance, diversity, and other population- or community-level characteristics of exposed invertebrate, fish, and wildlife communities (e.g., insectivorous songbirds and mink).

- Measuring reproductive success in fish (e.g., fecundity and hatching success in fish species such as largemouth bass and brown bullhead), birds (e.g., clutch size and hatching success for great blue heron and songbirds other than tree swallows), and mammals (e.g., uterine scar analysis and fetus counts from mink carcasses collected by New York State Department of Environmental Conservation [NYSDEC]).

- Measuring PCB concentrations in whole bodies or tissues of receptors other than fish; e.g., mink carcasses collected by NYSDEC, muskrat carcasses, great blue heron eggs, and songbird eggs other than tree swallow.

- Performing sediment toxicity test with survival, growth, and/or reproductive endpoints for species such as amphipods and chironomid larvae, which in combination with benthic macroinvertebrate surveys and sediment chemistry would have provided for a site-specific triad analysis.

Each type of measurement typically requires knowledge of, and data relevant to, the population dynamics of the species for appropriate use in assessing risks to populations. Measures of effects on individual organisms must be interpreted in the context of the distribution, abundance, and temporal dynamics of the exposed populations.

These methods are described in available EPA guidance documents and in the refereed scientific literature. Experience and practice at other comparable sites demonstrates the inadequacy of the BERA.

In its Responsiveness Summary, EPA claimed that the use of more site-specific studies for the BERA would have delayed the process several years. The reassessment was initiated more than 10 years ago. The question is not how long it would take to do an adequate risk assessment but whether this assessment fully passes muster.

The risk characterization should have relied on a weight of evidence (WOE) approach that evaluated endpoints and assigned differential weights to them based on the ecological relevance of each endpoint, the amount of site-specific data used to characterize each endpoint, and the relative uncertainty.

EPA claims (p. 167 of Hudson River BERA) to have used a "weight of evidence" approach to "assess the potential for adverse reproductive effects in the receptors of concern as a result of exposure to PCBs in the Hudson River." EPA's assessment, however, presents virtually no lines of evidence other than screening-level TQs and fails to present a framework for weighing individual lines of evidence and for resolving conflicting lines of evidence. Most of the assessment endpoints are addressed using only one line of evidence: comparison of measured or modeled exposure concentrations to generic, non-specific toxicity benchmarks, particularly Toxicity Reference Values (TRVs) and Sediment Effect Concentrations (SECs). Many of the so-called "lines of evidence" are based on the same or similar data and are not truly independent. For example, TQs for fish and wildlife are presented using TRVs based alternatively on total PCBs and TEQs. Both approaches to TRV-development are based on the same types of data; they simply have different theoretical foundations and use different exposure estimates. Similarly, the same water and sediment-quality-based TQs are cited as evidence for risks to every receptor group.

In the risk characterization for the Hudson River BERA, EPA weighted different lines of evidence equally and did not follow its own risk assessment guidance. This is not a simple process of counting up a number of studies and keeping score of the findings. The results of 10 bad studies cannot be compared on an equal footing with the results of one high quality, relevant study. The quality and relevance of each study must be closely evaluated.

The WOE method should have defined specific attributes for judging the quality of each measurement endpoint. Evaluation attributes are discussed in the *Guidelines for*

Ecological Risk Assessment (EPA 1998), and in the refereed scientific literature (Suter and Loar 1992; Suter 1993; Menzie et al. 1996; Suter et al. 1999). Examples of such attributes (from EPA 1998) that the risk assessor should consider when evaluating separate lines of evidence are:

- The relevance of evidence to the assessment endpoints.
- The relevance of evidence to the conceptual model.
- The sufficiency and quality of data and experimental designs used in key studies.
- The strength of cause/effect relationships.
- The relative uncertainties of each line of evidence and their direction.

In the WOE approach, empirical measurement endpoints, such as those suggested above, should be given greater weight than theoretical or modeled results. As EPA (1998) has said: "This process involves more than just listing the factors that support or refute the risk. The risk assessor should carefully examine each factor and evaluate its contribution to the risk assessment."

Given the size, complexity, and potential magnitude of remedial actions for the site, the BERA should have relied heavily on empirical site-specific data on the status of local populations and evidence for adverse effects of PCBs on those populations.

Many studies of the biological resources of the Hudson River and Valley have been carried out over the last 25 years. As a result, many sources of field data on the status of benthic communities and of fish and wildlife populations are available. These studies generally demonstrate the presence of healthy populations and communities in the upper and lower Hudson notwithstanding exposures to PCBs and other substances. EPA failed to use these data in the BERA.

Available data on the status of Hudson River biological resources conflict with EPA's conclusions in the Hudson River BERA. For example, the NYSDEC has examined macroinvertebrate populations in the Hudson River and could not identify any adverse effects from exposure to PCBs. National Marine Fisheries Service (NMFS) and the Fish and Wildlife Service (USFWS) analyzed the effects of PCBs on the striped bass population under the Atlantic Striped Bass Conservation Act and concluded that PCBs were not the cause of declines in the Hudson or coastal striped bass populations. NYSDEC has examined the growing number of eagles in the Hudson Valley and throughout the state. Breeding Bird Survey data document healthy populations of many other bird species in the Hudson Valley. These data and other similar data reflect the actual health of the wildlife populations of the Hudson River Valley.

EPA's failure to consider the extensive data available concerning the status and trends of fish populations in the lower Hudson is an unacceptable omission that clearly conflicts with agency guidance. In its Responsiveness Summary, EPA dismissed these data out of hand on the grounds that they "...are not indicative of biomass estimates in the river generally, but rather were designed to assess the impact of thermal discharges on fish populations," and on the grounds that other factors (e.g., fishing bans) might mask PCB-related effects. The first of these assertions is factually incorrect and the second is not supported by any data or analyses. The studies in question, which were recently documented in a Draft Environmental Impact Statement for the lower Hudson (Central Hudson Gas & Electric Corp., et al.1999), include studies performed both by the Hudson River utility companies and by federal agencies. They provide data on trends in riverwide abundance of striped bass, white perch, shortnose sturgeon (all of which are receptor species evaluated by EPA) and other fish species over the past 25 years. These data unequivocally demonstrate the existence of a healthy fish community in the lower Hudson.

The available field survey data should have been used in the BERA and weighted in the WOE approach in accordance with their value relative to data for other lines of evidence. Whether the observed increases in abundance of (for example) striped bass is masking adverse effects of PCBs can only be determined by a thorough evaluation of this line of evidence, in light of other evidence (e.g., time trends in recruitment or reproductive success vs. PCB exposure levels). EPA has performed no such evaluations. Obviously all factors affecting a fish population should be considered, but to claim that 25 years of comprehensive data is of no value in evaluating the affect, if any, of PCBs on fish populations amounts to the fallacious contention that no actual data from the population of issue are of any value.

Similarly, EPA does not address the apparent conflict between the results of TQ calculations and the empirical reproductive data available for bald eagles that nest in the Lower Hudson River. The reproductive status of the bald eagle population of the Lower Hudson River has been extensively investigated by NYSDEC. The results of these investigations indicate that the bald eagles are reproducing successfully. In three nests monitored since 1992 (NY18, NY28, and NY37), the average fledging success (number of young produced per occupied nest) has increased from 0 in 1996 to 0.5 in 1997 to 1.3 in 1998 to 1.7 in 1999. Sprunt et al. (1973) estimated that productivity of 0.7 young/occupied site was sufficient to maintain a stable bald eagle population (quoted by USFWS 1999). The recovery plans for bald eagle populations in the contiguous United States specify productivities of 0.9 to 1.1 young/occupied nest (USFWS 1999; Table 2). The average productivity of the resident bald eagle population of the Lower Hudson River from 1992 to 1999, 0.7 young/occupied site, is equal to the critical value of 0.7 estimated by Sprunt et al. (1973). The annual values for the past two years, 1.3 and 1.7, as well as the average value for the past three years, 1.25, exceed the recovery goals of the USFWS in the contiguous United States. In contrast, the TQ calculations for bald eagle suggest the potential for severe adverse reproductive effects. For example, the TEQ-based TQ for 1999 at mile 152 is greater than 100 (Table 5-41 of the BERA: average NOAEL, mile 152).

Although EPA mentions the empirical reproductive data in the BERA, and presents recent productivity data in the Responsiveness Summary, the conclusions that arise from this information were not integrated into the risk characterization. Their incorporation must involve resolution of an apparent contradiction: hazard quotients are high, up to more than 100, yet the bald eagles are reproducing successfully. Given this success, it is unclear to what degree reduction in a hazard quotient to near 1 would affect eagle productivity. The use of the bald eagle hazard quotients to drive management decisions must be considered in this light.

The only site-specific empirical data cited by EPA in the BERA (i.e., the benthic macroinvertebrate and tree swallow studies) do not demonstrate adverse effects of PCBs on the survival, growth, or reproduction of local populations. Moreover, GE is not aware of any empirical data that document these types of cause and effect relationships in Hudson River wildlife.

EPA's benthic macroinvertebrate study did not employ a design capable of separating effects of PCBs from effects of environmental variables such as water depth, grain size, total organic carbon (TOC), and presence of other potentially toxic chemicals. The results presented in Appendix H, Table H-6 of the BERA show that concentrations of PCBs, TOC, cadmium, chromium, lead, and mercury all co-vary at the five stations studied. Hence, although benthic community metrics differ between Stations 5 and 7 (higher PCB concentrations) vs. Stations 3, 4, and 6 (lower PCB concentrations), it is not possible to infer that PCBs are responsible for the differences in macroinvertebrate community metrics between these two groups of sites.

These results contradict the BERA's conclusion (pp. ES-6, 167) that PCBs are adversely affecting benthic macroinvertebrate populations in the Upper Hudson River. No effects of PCBs have been seen in Hudson River benthic macroinvertebrate communities as evidenced by the increase in abundance of pollution-intolerant filter feeders (NYSDEC 1993) over a 25 year period. Exponent (1998) studies for GE have found abundant and diverse populations of benthic macroinvertebrates, even in areas of the Upper Hudson River known to contain relatively high concentrations of PCBs in sediment.

In its Responsiveness Summary, EPA admitted the inconclusiveness of its macroinvertebrate study but defended its use in the BERA, stating that "The macroinvertebrate community study, in conjunction with sediment and water PCB concentrations, suggest an adverse effect of PCBs on benthic macroinvertebrates serving as a food source to local fish, which is consistent with the conclusion reached in the BERA for this endpoint." GE can only interpret this statement to mean that, even though the macroinvertebrate study was inconclusive, the results were used anyway because they could be interpreted to agree with the conclusions from the TQs. This reasoning is clearly inconsistent both with EPA guidance and with accepted principles of scientific inference.

In predicting effects on benthic macroinvertebrates, EPA's reliance on generic sediment quality values, unsupported by site-specific toxicity tests or defensible benthic community surveys, is inconsistent with best scientific practice in ecological risk assessment and fails to provide a sound basis for risk management decisions. The SECs for the Hudson River are simply another set of generic screening criteria that have no particular relevance to the issues specific to the Hudson River RI/FS and are not reliably predictive of ecological effects. Technical and conceptual flaws inherent in the approach used to derive the SECs result in sediment screening values inappropriate for use as indicators of the presence of or potential for adverse effects in benthic communities. These values can only be used to determine levels of PCBs in sediment below which adverse effects are unlikely. Any other use could result in an unacceptably high rate of false positives (as high as 40-50 percent). Moreover, the use of SECs does not address the assessment endpoint for this risk assessment, which concerns community-level effects.

Many researchers involved in the development of sediment quality values have cautioned that their use should be restricted to screening-level analyses to determine which sediments do not cause adverse effects (Long et al., 1995; Smith et al., 1996). The use of SECs to *predict biological effects* in a BERA is inconsistent with the intent of the sediment quality values as stated by their authors. O'Connor et al. (1998) provided the following cautions:

"*We conclude that guidelines based on bulk sediment chemistry can provide useful triggers for further analysis but should not be used alone as indicators of toxicity.*"

Most recently, Jones et al. (1999), in evaluating the impacts of heavy metals and PCBs on benthic invertebrate communities in the Clinch River and East Fork Poplar Creek, Tennessee, concluded that:

"*Sediment chemical concentrations indicated marginal or significant risks in all nonreference reaches. However, the community surveys and sediment toxicity tests often indicated that chemicals were either less bioavailable than those in sediments used to determine the benchmarks or the indigenous organisms were less sensitive to these chemicals than were the test organisms used to derive the benchmarks.*"

A baseline ecological risk assessment intended to support remedial action decisions for the Hudson River should employ site-specific evidence concerning benthic community effects and should not rely solely on bulk chemistry data and generic SECs.

Tree Swallows

McCarty and Secord have been unable to demonstrate an exposure-response relationship between PCB concentration and tree swallow reproduction. In fact, the reproductive performance of Hudson River tree swallows was actually higher at areas with higher PCB levels when compared to areas with lower PCB levels. The differences in reproductive parameters between the Ithaca (reference area) and Hudson River tree swallow populations are very likely due to the natural and temporal variation of these parameters between populations. The behavioral responses are not correlated with reproductive success.

In a study by Exponent for GE, data on tree swallow reproduction at various uncontaminated reference sites were compiled and compared with the Hudson River tree swallow data. For all reproductive parameters examined, the reproductive potential of the Hudson River tree swallows fell within the distribution of the reference area data (P>0.05 for a test of each value being outside the reference distribution). Further, the behavioral endpoints (which consist of nest quality metrics) mentioned in the BERA and measured by McCarty and Secord were not correlated with reproductive success. However, in its Responsiveness Summary, EPA uses abnormal nest building behavior as evidence of potential adverse reproductive effects in years of adverse weather or other adverse environmental conditions. This indicates that EPA considers speculative effects a stronger piece of evidence than empirical field data.

References

Central Hudson Gas & Electric Corp., Consolidated Edison Co. of New York, Inc., New York Power Authority, and Southern Energy New York. 1999. Draft Environmental Impact Statement for State Pollutant Discharge Elimination System Permits for Bowline Point, Indian Point 2&3, and Roseton Steam Electric Generating Stations. Contact: John Young, Consolidated Edison Co. of New York, 4 Irving Place, New York, NY 10003.

EPA. 1997. Ecological risk assessment guidance for Superfund: Process for designing and conducting ecological risk assessments. EPA/540/R-97/0CS. U.S. Environmental Protection Agency.

EPA. 1998. Guidelines for ecological risk assessment. EPA/630/R-95/002F. U.S. Environmental Protection Agency.

Exponent. 1998. Macroinvertebrate community and fish stomach content analysis for sampling conducted in 1997. Data report. Exponent, Bellevue, WA.

Jones, D.S., L.W. Barnthouse, G.W. Suter II, R.A. Effroymson, J.M. Field, and J.J. Beachamp. 1999. Ecological risk assessment in large river-reservoirs. 3-Benthic Invertebrates. Environ. Toxicol. Chem. 18(4):599–609.

Long, E.R., D.D. MacDonald, S.L. Smith, and F.D. Calder. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. Environ. Manage. 19:81-97.

Menzie, C., M.H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts weight-of-evidence workgroup: a weight-of-evidence approach for evaluating ecological risks. Human Ecol. Risk Assess. 2(2): 277–304.

NYSDEC. 1993. 20 year tends in water quality of rivers and streams in New York state based on macroinvertebrate data, 1972–1992. New York State Department of Environmental Conservation, Albany, NY.

O'Connor, T.P., K.D. Daskalakis, J.L. Hyland, J.F. Paul, and J.K. Summers. 1998. Comparisons of sediment toxicity with predictions based on chemical guidelines. Environ. Toxicol. Chem. 17(3):468-471.

Smith SL, MacDonald DD, Keenleyside CG, Ingersoll CG, Field J. 1996. A preliminary evaluation of sediment quality assessment values for freshwater ecosystems. J. Great Lakes Res. 22:624-638.

Sprunt, A., IV, W.B. Robertson, Jr., S. Postupalsky, R.J. Hensel, C.E. Knoder, and F.J. Ligas. 1973. Comparative productivity of six bald eagle populations. In *Transactions of the 38th North American Wildlife and Natural Resources Conference*, pp. 98-1-6.

Suter, G.W. II (ed.). 1993. Ecological risk assessment. Lewis Publishers, Chelsea, MI.

Suter, G.W. II, and J. M. Loar. 1992. Weighing the ecological risks of hazardous waste sites: the Oak Ridge case. Environ. Toxicol. Chem. 26:432–438.

Suter, G.W. II, L.W. Barnthouse, R.A. Effroymson, and H. Jager. 1999. Ecological risk assessment in a large river reservoir: 2 fish community. Environ. Toxicol. Chem. 18(4):589–598.

U.S. Fish and Wildlife Service. 1999. Endangered and Threatened Wildlife and Plants; Proposed Rule to Remove the Bald Eagle in the Lower 48 States From the List of Endangered and Threatened Wildlife; Proposed Rule. Federal Register 50 CFR Part 17, July 6, 1999.

Appropriateness of TRVs

The baseline ecological risk assessment (BERA) peer review panel was asked to comment on the validity of developing only a no-observed-adverse-effect level (NOAEL) toxicity reference value (TRV) for field-based studies and to comment on the general approach of using uncertainty factors in developing TRVs (Specific Question 6).

In developing its risk conclusion, EPA's primary and most heavily weighted line of evidence is toxicity quotients (TQs). The TRVs used in the TQ calculations were developed using a conservative approach designed primarily for screening assessments. The excessive conservatism EPA has built into the TRVs results in highly biased and overstated risk conclusions. While this approach may be appropriate for a screening risk level assessment, it is inappropriate for a quantitative baseline risk assessment that is intended to support a remedial decision for a complex site such as the Hudson River.

Validity of Developing NOAELs From Field Studies

EPA calculated TQ values from field-study derived NOAEL values, and interpreted exceedances of these values the same way exceedances of laboratory-study derived, bounded NOAEL or lowest-observed-adverse-effect level (LOAEL) values were interpreted. In fact, field-derived NOAELs were given precedence if they were derived for species in the same taxonomic family as receptor species. This approach significantly understates the uncertainty associated with field-derived NOAELs.

The use of field studies to develop TRVs requires considerable caution. Well-designed laboratory toxicity tests produce estimates of both a NOAEL, the highest exposure at which no adverse effect is observed, and a LOAEL, the lowest exposure at which an adverse effect is observed. The NOAEL from such a study is usually assumed to represent a threshold above which toxicity may occur in exposed organisms. Results derived from field studies are much more difficult to interpret. Numerous investigators have published PCB concentrations in tissue (e.g., whole body, muscle, egg, liver, brain) both from populations of organisms that do not exhibit adverse effects and from populations that do exhibit effects. Most of these studies are conducted in areas contaminated with multiple additional compounds such as DDT and PCDDs and PCDFs. If no adverse effects are observed, the measured PCB concentrations can be considered to be NOAELs. However, if adverse effects *are* observed, they may have been caused by chemicals other than PCBs. A LOAEL for PCB exposure cannot be defined. A NOAEL derived from a study in which no LOAEL can be determined is termed an "unbounded" NOAEL. Unbounded NOAELs cannot be assumed to represent toxicity thresholds. The true threshold of toxicity is unknown, and may be far more than 10 times higher than the NOAEL. Therefore, exceedance of an unbounded NOAEL does not imply that adverse effects are likely to occur. Because the magnitude of an exceedance of an unbounded NOAEL is not necessarily a measure of risk, exceedances of field-study derived NOAEL values should only be expressed qualitatively (e.g., as pass or fail) and as such, are more appropriate for use in a screening-level assessment.

Appropriateness of Using Uncertainty Factors in Developing TRVs

The method EPA uses to develop TRVs for the baseline ecological risk assessment (BERA) (p. 79–81) involves selection of a single laboratory or field study (usually the lowest available number), and uncritical application of generic uncertainty factors of 10 for NOAEL-to-LOAEL, subchronic-to-chronic, and interspecies extrapolations. These conservative safety factors are multiplicative, resulting in TRVs that are, in many cases, many times lower than the lowest concentration or dose ever observed to affect exposed organisms. This approach does not represent best or current scientific practice and is unreasonable, especially considering the weight EPA has placed on the TQ calculations to support the risk conclusions.

EPA applies an uncertainty factor of 10 to convert LOAELs to NOAELs, and does not provide any specific scientific support for the use of this factor. Chapman et al. (1998) recently reviewed the use of uncertainty factors to extrapolate from LOAELs to NOAELs in risk assessment, and concluded that the LOAEL-to-NOAEL uncertainty factor "should not be used in ecological risk assessment." Chapman et al. (1998) further states that "[t]o estimate a NOEC from a LOEC by merely dividing the LOEC by 10 is compounding the uncertainty in a manner that makes the result essentially meaningless." Human health and ecological risk assessors recognize the limitations associated with this approach and have begun "moving away" from the use of this approach "in favor of a benchmark value," namely the confidence interval of the concentration that results in an certain level of response (Chapman et al. 1998).

EPA applies another uncertainty factor of 10 to adjust NOAEL values to extrapolate between species if the test and receptor species are not within the same taxonomic family, unless the test species is known to be the most sensitive of all species tested. EPA refers to methodologies developed for assessment of cancer risks in humans in support of this method. For noncancer risks, however, participants in an EPA-sponsored workshop clearly concluded that default factors of 10 are generally excessive (Abdel-Rahman 1995). Moreover, there is a clear philosophical difference between the human health risk assessment and ecological risk assessment. The basis for human health risk assessment is protection of all humans as individuals, whereas the basis for ecological risk assessment is the protection of ecosystems. The protection of individuals may warrant greater conservatism.

EPA applies yet another uncertainty factor of 10 to estimate chronic NOAELs from subchronic NOAELs and cites "recently developed guidance" by Oak Ridge National Laboratory (Sample et al. 1996). Sample et al. (1996) does not provide recommendations or represent guidance on the use of uncertainty factors. Rather, Sample et al. (1996) presents toxicological benchmarks for assessment of certain chemicals on mammalian and avian species. Moreover, Sample et al. (1996) considers studies to be representative of chronic effects if the exposure occurs during critical life stages (e.g., reproductive phase), regardless of exposure duration. Therefore, Sample et al. (1996) does not use a factor of 10 to convert subchronic NOAEL and LOAEL values to chronic values if the exposure occurs during a reproductive phase of the test organism. In contrast, EPA consistently decreases subchronic toxicity thresholds by a factor of 10, regardless of the time of exposure.

In the recent review of the use of uncertainty factors for ecological risk assessment, Chapman et al. (1998) made the following relevant conclusions and recommendations:

> Values derived using safety factors should not be used as threshold values for a toxic effect or as absolute values.
>
> Uncertainty, by definition . . . , is an integral component of risk assessment and is usually addressed by conservatism (e.g., the use of large safety factors). Probabilistic modeling is a better method of addressing uncertainty. However, the best method to address uncertainty is to measure the same endpoint using different approaches and assign the greatest credibility to results that are confirmed by a combined evaluation of the approaches used (i.e., weight of evidence . . .). For example, when chemical analyses indicate a potential problem but appropriate biological studies indicate no problem, corrective action would not generally be necessary.
>
> Safety factors are not intended as mathematical absolutes but rather as screening tools that are surrounded by some unquantifiable level of imprecision. Such screening tools may be suitable for use by trained risk assessors, but they can be deceiving and confusing to the general public, particularly when there are disagreements of an order of magnitude or more between different numbers.
>
> Extrapolation requires context. Any use of safety factors should be based on existing scientific knowledge and should include appropriate caveats.
>
> Extrapolation is not fact. Safety factors should be used only for screening, not as threshold or absolute values.
>
> Extrapolation is uncertain. Safety factors should encompass a range, not a single number.

Thus, generic uncertainty factors or safety factors of 10 are an overly conservative and imprecise approach for dealing with uncertainty associated with assessing chemical risks.

The uncertainty factor approach used by EPA involves unscientific and unqualified application of arbitrary multiplicative factors that have no known relevance to actual uncertainty. Although application of this type of approach may reduce the probability of underestimating risk, it greatly increases the probability of overestimating risk, and leads to unrealistic and unscientific conclusions in the risk assessment. The uncritical manner in which EPA has used these excessive safety factors thus represents a policy decision, not a science-based decision. Their use does not support an accurate quantitative risk characterization, and therefore does not support the development of reasonable remedial alternatives. In effect, the risk manager is faced with a highly uncertain risk assessment in which the TQs have little or no scientific meaning. This is not a reliable basis for making major risk assessment decisions.

References

Abdel-Rahman, M.S. (ed). EPA uncertainty factor workshop. Hum. Ecol. Risk Assess. 1:512–662.

Chapman, P.M., A. Fairbrother, and D. Brown. 1998. Annual review: a critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environ. Toxicol. Chem. 17(1):99–108.

Sample, B.E., D.M. Opresko, and G.W. Suter, II. 1996. Toxicological benchmarks for wildlife: 1996 revision. ES/ER/TM-86/RS. Prepared for the U.S. Department of Energy, Office of Environmental Management. Oak Ridge National Laboratory, Risk Assessment Program, Health Sciences Research Division, Oak Ridge, TN.