

**Toxics Release Inventory  
Relative Risk-Based Environmental Indicators:**

**Summary of Comments Received on the  
Draft 1992 Methodology  
and Responses to Comment**

Nicolaas W. Bouwes, Ph.D.  
Steven M. Hassur, Ph.D.

Economics, Exposure and Technology Division  
Office of Pollution Prevention and Toxics  
U.S. Environmental Protection Agency

May 1997

Contractor Support:

Abt Associates, Inc.  
4800 Montgomery Lane  
Bethesda, MD 20814

For further information or inquiries, please contact:

Nicolaas W. Bouwes, Ph.D.

(202) 260-1622  
bouwes.nick@epamail.epa.gov

or

Steven M. Hassur, Ph.D.

(202) 260-1735  
hassur.steven@epamail.epa.gov

Economics, Exposure and Technology Division (7406)  
Office of Pollution Prevention and Toxics  
U.S. Environmental Protection Agency  
401 M St., SW  
Washington, D.C. 20460

## Table of Contents

I.	BACKGROUND .....	1
II.	KEY TO COMMENTS .....	3
III.	GENERAL COMMENTS ON THE TRI INDICATORS METHODOLOGY .....	4
	Issue: Need for the TRI Environmental Indicators .....	4
	Issue: Adequacy of TRI Indicators for Tracking Environmental Impacts .....	4
	Issue: Methodology Gives Appearance of Risk Assessment .....	7
	Issue: Complexity of the TRI Environmental Indicators .....	8
	Issue: Use of Conservative Values will Magnify Impacts .....	9
	Issue: Use of Single Versus Several Indicators .....	9
	Issue: Recommendations for Further Review and QA/QC of the Indicators Methodology .....	11
	Issue: Reporting and Interpretation of the TRI Environmental Indicators Results .....	14
	Issue: Public Perception of the TRI Environmental Indicators .....	18
	Issue: Acute Effects Indicators .....	20
	Issue: Expand Methodology Beyond TRI Reporting .....	22
	Issue: Selection of Chemicals and Industries .....	23
IV.	TOXICITY WEIGHTING .....	28
	Issue: Toxicity Ranking .....	28
	Issue: Basing Weight on Most Sensitive Endpoint .....	30
	Issue: Method Does Not Weight for Severity of Effect .....	32
	Issue: Toxicity Data Used for Weighting .....	33
	Issue: Weight of Evidence Discussion .....	37
	Issue: Other Related Comments .....	40
V.	TRI CHRONIC ECOLOGICAL INDICATOR .....	41
	Issue: Indicators Do Not Address Terrestrial or Non-Water Column Wild Populations .....	41
	Issue: Ecological Impacts .....	42
	Issue: Ecological Toxicity Data .....	43
	Issue: Ecological Toxicity Weighting Approach .....	44
	Issue: Calculation of Bioconcentration Factors .....	48
	Issue: Weight of Evidence Considerations .....	48
	Issue: Ecological Exposure Weighting .....	49
	Issue: Calculation of Ecological Indicator .....	49
VI.	OTHER ISSUES .....	51
	Issue: Technical Questions and Corrections About Use of TRI Data .....	51
	Issue: Environmental Modeling Assumptions/Approaches .....	52
	Issue: Aggregate Population Risk (Versus MEI Risk) .....	60

Issue: TRI Environmental Indicators and GIS .....	62
Issue: Calculation of the TRI Environmental Indicators .....	63
Issue: Normalizing the TRI Environmental Indicators .....	63
Issue: Expansion of the TRI Environmental Indicators .....	64
Appendix .....	A-1

## I. BACKGROUND

The Office of Pollution Prevention and Toxics (OPPT) has developed an environmental indicators methodology, based on the Toxics Release Inventory (TRI), to assess the releases of TRI and other chemicals from a relative risk-based perspective. The TRI Relative Risk-Based Chronic Human Health Indicator provides a risk-related measure of the relative impacts of chemical emissions regarding chronic human health.

Four Indicators are eventually planned (chronic/acute human and ecological). The Indicators are numeric relative ranking values, based upon reported TRI multimedia emissions and weighting factors representing toxicity, exposure characteristics, and receptor populations based upon current EPA models and databases. The Indicators use TRI data and models to calculate Indicator Elements for each combination of facility, chemical, and media reported under TRI. Each Indicator Element reflects a surrogate dose weighted by toxicity and exposed receptor population. Each year of TRI reporting generates approximately 400,000 of these Indicator Elements; which when summed form the Indicator. By comparing each year's Indicator to the base year of 1988, or to a previous year, one can obtain a risk-based perspective of trends in environmental well-being as a function of chronic human health. In addition to relative risk-based trends analysis, the Indicators can also be utilized for targeting and prioritization. This tool can effectively conserve Agency resources in project planning and analysis; it also has environmental justice applications.

The Indicators methodology will be submitted for formal review by the EPA Science Advisory Board on July 2, 1997. A complete discussion of the methodology is available to the public in the document, *Toxics Release Inventory Relative Risk-Based Environmental Indicators Methodology* (May 1997).

The present document provides a detailed summary of comments received from 25 reviewers of the Draft Toxics Release Inventory (TRI) Environmental Indicators Methodology document (May 1992 and September 1992 drafts). The complete set of comments from the three peer reviewers are attached as an appendix. This summary is organized by topic or issue area and the comments (keyed by individual commenter) are grouped in the following manner: Peer Reviewers, Agency Headquarters (HQ), Agency (Regions), Industry, Environmental Groups, where each category of commenter is appropriate. Associated with each issue or specific comment(s) is a response reflecting the view of project staff, and changes that were made in either the Indicators methodology or in the description of the Indicators approach. There was a diversity of opinion expressed regarding certain issues and complete consensus was not always achieved.

The May 1992 methodology document which was reviewed by the three outside peer reviewers reflected our initial desire to produce a single, national indicator (with no opportunity for users to manipulate or examine the underlying data). The Indicators originally tried to mimic the *Dow Jones Industrial Average*, by attempting to evaluate a subset of representative chemicals and, to some extent, industry sectors.

Based on the results of this peer review, some changes were made in the September 1992 methodology document which was made available to the public for general review (concurrent

with a notice in the Federal Register and a public meeting in Chicago, IL). Partly in response to the peer reviewer comments, the Indicators methodology was modified to reflect more of the analytical approach desired by these three reviewers, and to address all TRI chemicals which have sufficient information for inclusion. However, the methodology still does not correspond to the site-specific, quantitative risk assessment desired by some, since this is not OPPT's purpose in constructing these Indicators as a relative risk-based, screening-level tool.

## II. KEY TO COMMENTS

The reference to the position held by each individual in the key below, reflects the mid- to late-1992 timeframe of these comments.

### Peer Reviewer Key:

- (a) Adam M. Finkel, Sc.D., Resources for the Future
- (b) D. Warner North, Ph.D., Decision Focus Incorporated
- (c) John D. Graham, Ph.D., Harvard Center for Risk Analysis

### Agency Headquarters (HQ) Key:

- (d) Margaret Stasikowski, Health and Ecological Criteria Division, Office of Water
- (e) Kim Devonald, Environmental Results Branch, OPPE
- (f) Robin Heisler, Source Assessment and Information Management Branch, Groundwater Protection Division, Office of Ground Water and Drinking Water
- (g) Richard Wilson, Office of Mobile Sources, Office of Air and Radiation
- (h) Sylvia Lowrance, Office of Solid Waste
- (i) William Silagi, Regulatory Impacts Branch, ETD, OPPT (2 memos)
- (j) Joe Cotruvo, Chemical Screening and Risk Assessment Division, OPPT
- (k) Vanessa Vu, Health Effects Branch, HERD, OPPT
- (l) Maurice Zeeman, Environmental Effects Branch, HERD, OPPT

### Agency Regions Key:

- (m) William Patton, Pesticides and Toxic Substances Branch, Region IV
- (n) Lee Gorsky, Assistant Regional Health Advisor, Region V
- (o) William Sanders, Environmental Sciences Division, Region V
- (p) A. Stanley Meiburg, Air, Pesticides and Toxics Division, Region VI
- (q) Robert Jackson, Toxic Substances Control Section, Region VII
- (r) C. Alvin Yorke, Toxic Substances Branch, Region VIII

### Industry Key:

- (s) Cheryl Sutterfield, Pharmaceuticals Manufacturers Association
- (t) Lorraine Cancro, Hoffmann-La Roche
- (u) Joe Mayhew, Chemical Manufacturers Association
- (v) Walter Quanstrom, Amoco
- (w) D.J. Smukowski, Boeing

### Environmental Groups Key:

- (x) Working Group, Community Right to Know
- (y) Unison Institute

### **III. GENERAL COMMENTS ON THE TRI INDICATORS METHODOLOGY**

#### **Issue: Need for the TRI Environmental Indicators**

##### **Industry comments:**

We acknowledge the need for a program like Environmental Indicators and support the Agency's initiative in this regard. We believe that without a system to benchmark and measure environmental progress, scarce resources may be misdirected into activities which will not achieve the greatest benefits for the most people. This program, if properly planned and executed, can provide the knowledge base necessary to determine the need for, and prioritization of, future Agency actions. (v)

The public and EPA already have access to a wide variety of information. We hope that your committee will fully utilize existing data, on a national or regional basis, before developing new sources. Current sources of information, such as the SARA 313 Pollution Prevention Reports, have not been fully utilized. In addition, the need for new indicators should be weighed against the confusion that could result from another influx of SARA 313 data. (w)

##### **Agency (HQ) comments:**

In describing what the "ultimate goal of the indicator effort..." there is no mention that it should be meaningful. Although it may be inferred, I think commitment to a meaningful system of indicators should be clearly stated. (l)

*Response to above comments: Text acknowledges the importance of producing Indicators with meaningful (i.e., risk-related) results. The Indicators are based upon current TRI reporting requirements.*

#### **Issue: Adequacy of TRI Indicators for Tracking Environmental Impacts**

##### **Peer Reviewers comments:**

...system could be valuable in support of strategic planning and the Agency's communication to the public, in support of "the Agency's desire to set priorities and shift resources to areas with the greatest opportunity to achieve health risk and environmental risk reductions." (p. 1 of your draft). You recognize the limitations of the TRI and other data sources that reflect multimedia trends in environmental contaminant releases. The methodology must build upon the data available. Unfortunately, this is a very large limitation because of the lack of data and unevenness in data quality. In particular, you face enormous difficulties in trying to assess what emissions imply for exposure to humans and environmental receptors. Without models and the data to drive them, source-to-receptor relationships can be described only in the most generic terms. Further, chemical substances transform into other substances that may be more toxic, less toxic, or non-toxic. How quickly such transformations occur may depend on specific characteristics of the medium - air, water, soil -that can vary in time and space and with the concentrations of other chemicals present. So even with very good data on emissions, it can

be very difficult to assess the consequences of the emissions in terms of risk to health and the environment. (b)

*Response: OPPT agrees that site-specific and chemical-specific risk assessment is difficult, especially when lacking the highest quality emissions data. However, OPPT feels that this screening-level tool does utilize the available data in a reasonable manner to estimate relative risk-based impacts. It uses currently available Agency databases and models as a basis for its calculations, but is not designed for quantitative risk assessment.*

### **Environmental Groups comments:**

We appreciate the tremendous effort that has gone into development of this document, however, we feel the Agency's resources could be better directed. The limitations of the TRI Environmental Indicator methodology and of the indicators themselves as a measure of facility improvements make indicator development wholly inadequate for evaluating success in reducing environmental risks. These limitations (which we will emphasize to both the press and the public) include:

- Environmental indicators should not be based on Toxics Release Inventory (TRI) data using this methodology. TRI data do not account for exposures to toxic chemicals by workers, consumers and populations affected by post-consumer recycling, treatment, and disposal of toxic household products. Additionally, TRI data do not cover potential hazards to communities from transportation of toxic chemicals. We believe the preferred data input for environmental indicators is toxic chemical use data (as would be collected under proposed federal Community Right-to-Know More legislation), not TRI data. For these reasons, US EPA should support Congressionally-mandated collection of data on toxic chemical use.
- TRI Environmental Indicators could show improvement even when reporting facilities have not made any actual changes. For example, should nearby populations decrease, TRI Environmental Indicator scores would automatically decrease, even though the environmental risk to individuals who still live near such facilities would be unchanged. Along these lines, the methodology also is unable to tell the public anything about highly exposed individuals.
- The methodology does not account for chemical synergism resulting in increased hazards to human or ecosystem populations. Even if the methodology could be modified to account for synergism, toxicity data on multiple chemical exposures generally does not exist.

In light of these limitations, we believe that EPA should redirect its limited resources to: enhancing TRI data quality, expanding TRI (i.e., to include additional chemicals and facilities, and collection of peak release data), and advocating for new authority to collect toxic chemical use data to provide a credible basis for pollution prevention programs. (x)

We agree that the first step (toxicity ranking) [in the methodology] is valuable, disagree that the second step [exposure ranking] is useful because of its undue complexity and incorrect

assumptions, and believe that the third step [population weighting] produces a result which can be dangerously misused. We will go into our concerns with each step in detail. Because of these concerns we have three major recommendations for EPA:

- Shift the focus of the indicators from relative risk to pollution prevention by de-emphasizing the exposure and population calculations;
- Simplifying the exposure calculations wherever possible, in particular switching from modeling to assignment of constant, facility-independent numerical factors. This will allow release of the toxicity rankings and exposure factors as simple tables of numbers which can be used by others independently of the "Indicators" project;
- Re-examine some of the assumptions in the exposure calculations, such as the assumptions of zero exposure from underground injection wells, hazardous waste landfills, and hazardous waste incinerators. (y)

We would encourage EPA and OPPTS in particular to move from a set of risk rankings to a set of pollution production rankings based on the pounds of waste generated, released, transferred, and its toxicity. This would fulfill the goal of having a TRI Environmental Indicator while incorporating OPPT's pollution prevention mission and remaining sensitive to environmental equities issues. (y)

*Response to above comments: The Indicators are based within the framework of data gathered through existing reporting regulations (e.g., TRI reporting). The Indicators bring an important new risk-based perspective to the interpretation of TRI information. TRI reporting data does not include, workers, consumers, etc; but it has been expanded to include a much larger set of chemicals and chemical categories, as well as a larger universe of facilities. As a result, the Indicators are restricted to observing relative risk-based impacts on the general population. The method focuses on general population risk (i.e., cumulative risk vs. individual risk) and larger overall risks may be associated with exposure to large local populations. Presently, the project staff is not aware of a suitable methodology which addresses synergistic or antagonistic effects on large numbers of chemicals. Since OPPTS policy emphasizes prioritization and analysis from a risk-based perspective where possible, modeling of exposure to populations beyond the facility fence line is considered an important feature of the Indicators. Underground injection is no longer ignored by the Indicators (users may investigate the pounds of such disposal). The methodology discusses this in greater detail and proposes an alternative strategy.*

## **Issue: Methodology Gives Appearance of Risk Assessment**

### **Agency (HQ) comments:**

You should give added emphasis to the 1st sentence [that the method is not risk assessment; it will sell better to suspicious types. (l)

*Response: Additional caveats were added to the report to state that the indicator does not calculate exposure and cannot be interpreted as risk assessment. However, those who are uncomfortable with the approach are unlikely to be reassured by additional disclaimers.*

### **Agency (Regions) comments:**

The commentator is very concerned that, in spite of various disclaimers in the document, the methodology proposed is a risk assessment, since it follows the method by which risk assessments are performed. The TRI Indicator proposal seems to argue that risk assessment utilizes "precise measured values," but many risk assessments are carried using estimating procedures. The emissions information from TRI are not useable for exposure calculations and cannot be massaged into a form that would make them useable. TRI data contain no information about the time sequence of release. Without more information, it is not possible to make a credible estimate of exposure that is any better than a guess. (o)

Given the limitations of the existing TRI data, it is not appropriate to use TRI data in any methodology that calculates exposure. Thus, TRI data cannot be used in any methodology that approximates risk assessment. (o)

*Response: It is true that the issue of timing of release is relevant to exposure assessment. However, OPPT has decided that it is appropriate to use TRI data to gauge relative exposure potential, in the context of an indicator and screening-level prioritization and targeting tool. The Agency continues to perform data quality surveys to improve estimation and reporting of TRI emissions. Additional caveats have been added to the methodology to state that the Chronic Human Health Indicator does not calculate exposure and cannot be interpreted as risk assessment. The output of the model is a unitless relative ranking of risk-related impacts, not a quantitative risk estimate.*

The goal of tracking progress of the TRI reporting program is a good one. Likewise, identifying levels of concern over chemicals that may be carcinogens, developmental toxicants, neurotoxicants, or cause other health end points is useful for setting priorities for action and answering questions about exposure. Nevertheless, the methodology used in this document is fatally flawed. The substitution of numerical values for cancer, mutation, developmental toxicity, neurotoxicity and other weight of evidence schemes, and the use of these numerical values as a surrogate for risk assessment is a misuse of the process. We recommend formal review of this document by scientists with risk assessment training. (r)

This methodology should not be used as an indication of risk. The estimations of the toxicological significance of TRI chemicals cannot be construed to be surrogates for risk because, as noted on page 2, they do not include site specifics of exposure and target populations other than population numbers. In addition, the draft quantification of weight-of-evidence schemes imply for example, that release of a chemical with limited evidence of genotoxicity, including mixed positive and negative results (category 2-4) is 2-4 times less hazardous than one with evidence of genotoxicity in non-mammalian germ cell assays (category 8). Along the same lines, release of an "A" carcinogen (category 8-9) represents 1.3 times the risk of an equal amount of a B1 carcinogen (category 6-7). If the progress index were used to show how much risk had been reduced, it would be highly misleading. Further, if the index evolved into a funding, research, priority-setting or enforcement targeting mechanism, releases representing significant risks might be overlooked. (r)

*Response: These comments are factually incorrect - the method has never used the weight of evidence category numbers in the calculation - the numbers only identify categories. The numbers could just as well be letters or symbols. Formal review of the Indicators methodology by EPA's Science Advisory Board (SAB) is planned on July 2, 1997.*

### **Issue: Complexity of the TRI Environmental Indicators**

#### **Agency (HQ) comments:**

The Indicator may be too ambitious due to complexity of matrix of rankings. (d)

We feel the index should return to its original intent and be kept as simple as possible, and rely as much as possible on hard data and Agency reviewed assessments, and avoid extensive modeling beyond basic environmental distribution. It is essential that the index be understandable and credible to its users. As we gain experience and public acceptance in use of the basic system, the methodology can become more elaborate as data permit. (j)

The model and process should not become overly complex. Since the model is to be used for ranking and priority purposes, it should remain simple. (j)

The calculation section is confusing (scoring). Calculations should be made simple and the results easy to communicate to the public. (j)

*Response to above comments: The Indicators are based upon existing EPA models that have been used to communicate risk to the public. The Indicators can be aggregated at the national level and disaggregated at the geographic, chemical and facility level as well. This requires modeling, but the results can be made usable for a wide audience. OPPT believes that the public is capable of understanding risk-based conclusions generated by the Indicators, as easily as simpler models which require an understanding of what is omitted to have an appreciation for their results.*

### **Industry comments:**

We are concerned that the project is far too massive and ambitious to have significant use or relevance. The level of effort required to put this model together and to run it is very large relative to the value of the output gained from it. (u)

*Response: The Indicators represent a unique and valuable approach to relative risk-based analysis.*

### **Issue: Use of Conservative Values will Magnify Impacts**

#### **Industry comments:**

One concern is that conservative values will be developed for individual impacts which will then be multiplied together with the resultant value having an implied impact so highly magnified that the indicator will be a meaningless value. Observing environmental release trends is a complex and tremendous task, considering that so many variables need to be accounted for, so whatever unit is developed for measuring such a trend needs to be as accurate as possible without inflated factoring. Indicators derived using inflated factors may in fact suppress the trends they were intended to reflect. (t)

*Response: The commenter appears to be worrying about the exclusive use of unrealistic, upper bound estimates. The Indicators methodology uses a mixture of central tendency estimates, such as average body weight, a population weighted average vs. most exposed individual, etc., as well as high-end estimates. The method is directed at an overall comparative analysis and, in this relative sense, the high-end estimates may tend to cancel.*

### **Issue: Use of Single Versus Several Indicators**

#### **Peer Reviewers comments:**

You go from "indicators," plural, as what the Administrator wanted in paragraph 1 to the singular, "indicator," in paragraph 2. As discussed above, I think this change in emphasis to a single indicator is a poor idea. The discussion seems to stay with the singular from then on: "the methodology and data sources used to develop the TRI indicator" (page 1, last two lines). In practice, you should be using at least two indicators, for health and for environmental effects. (b)

*Response: The language has been changed to reflect that the method presents a Chronic Human Health Indicator and a Chronic Ecological Indicator.*

My basic concern with the draft is the quest for a single indicator. While the indicator methodology is presented as a risk-based analytical tool, the aggregation necessary in the computation requires numerous debatable assumptions which are outside the province of science. In my opinion, these assumptions are more appropriately made by risk managers, not risk assessors. The draft acknowledges this point to some extent by breaking down the summary

indicator into four components: acute human effects, chronic human effects, acute ecological effects and chronic ecological effects... Since you intend to design a methodology for use on a computer, I recommend that you move toward designing a data base rather than a set of summary numbers. Stated differently, EPA should be in the business of producing an array of summary indicators rather than a single, aggregated indicator. The data base would allow different users (with potentially different interests) to construct their own summary indicators based on their preferred weights and aggregation techniques....Given the large number of dimensions of environmental quality, EPA is simply asking for trouble if it chooses to produce and report only one summary indicator of environmental quality. By selecting a single indicator that rests on numerous uncertain assumptions, EPA is setting itself up for criticism and embarrassment in the future... Note that this concern can be addressed by simply recasting the draft report as a project in database design rather than as a project to develop a TRI-based indicator methodology. By reporting several indicators rather than one summary indicator, EPA would also foster greater public understanding of the multi-dimensional nature of environmental quality...If such a data base were available, consider how the public might use it. One user might be interested in the total number of people exposed to any chemical thought to be a possible, probable, or known carcinogen. Another user might be interested in the number of people exposed to chemicals that are known human carcinogens. One user might want to select a different baseline than another. Users might differ in what chemicals they want to include or exclude...It is true that such a data base would be difficult to use, requiring the user to understand the inputs, weights and aggregation options. However, the proposed TRI indicator is easy to use only because it does not require an understanding of its methodology. It is therefore difficult to imagine that its output will be wisely used. Any user who does fully understand the output of the proposed system has all the knowledge necessary to set the parameters in the data base. More importantly, any user who has the necessary knowledge probably would want to construct his or her own summary indicator(s) because users will be faced with different risk management questions and concerns. (c)

*Response: Although ideally OPPT would want both acute and chronic indicators, the TRI data do not support this. Therefore, the current focus is on only two chronic indicators - chronic human health and chronic ecological. While it is still possible to have a single summary Indicator for each of these two, the computer model and the associated database it generates can be used to aggregate and disaggregate Indicator Elements in order to address different types of questions or subsets of data a user may wish to address. In designing this methodology, the Indicators have benefited from the input of both risk managers and risk assessors.*

Your discussion of the current TRI data base should acknowledge that total mass emissions (aggregated over media and sources, for example) is itself a summary indicator based (implicitly) on various assumptions (e.g., each chemical emission is equally bad). By acknowledging that mass emissions is a summary indicator and reporting it next to other summary indicators, insight will be provided into the sensitive weighting and aggregation issues that you have identified. (c)

*Response: The current computer algorithm allows users to compute both a pounds-only Indicator and a full-method Indicator for these types of comparisons. Other options available to the user include an examination of pounds of emissions modeled by the Indicator, modeled pounds times toxicity, and modeled pounds times toxicity time population.*

**Agency (HQ) comments:**

Changes to the methodology that are aimed at reporting additional information could result in over-complicating the indicator. While the additional information would be interesting, perhaps the best solution is to wait until some actual test data resulting from using the indicator have been evaluated before making any additions. (h)

**Issue: Recommendations for Further Review and QA/QC of the Indicators Methodology**

**Peer Reviewers comments:**

As with risk estimation in general, I firmly believe that estimates such as "indicators" can be reduced to single point values, but only if these are conscious and consistent choices emerging from a quantitative uncertainty analysis (QUA) rather than from a black box containing a mixture of "average," "conservative," and "anti-conservative" procedures. I realize that a quantitative calculation of uncertainty may be daunting if done routinely, but I urge you to urge EPA to at least undertake state-of-the-art QUAs for a few illustrative cases within the TRI project. That way, an appreciation for the "noise" in the indicator can be communicated. More importantly, regulators and the public could be educated regarding which comparisons (e.g., longitudinal comparison of indicators over time, snapshots across regions, industrial sectors, etc.) involve the compounding of uncertainties and which probably involve the canceling of uncertainties. For example, suppose the indicator moves from "100 ± 20" to "80 ± 15" between 1995 and 1997. If the uncertainties are not strictly parallel, then I would argue the "noise" outweighs the "signal" of progress (the 1995 value might well have been 85, and the 1997 value might well have risen to 90). However, if a model QUA revealed that the sources of error were common to both measures in this comparison, then the apparent rank order might also be robust. (a)

*Response: Quantitative uncertainty analyses are planned for the Indicators model. Initial attempts to obtain relevant risk assessment data from a wide variety of sources found a few air-related studies (candidate risk assessments were collected from EPA's Office of Policy Planning and Evaluation, Office of Pesticides and Toxic Substances, Office of Solid Waste, the New Jersey Department of Environmental Protection, and the California Environmental Protection Agency) which address quantitative risks to only the highest exposed individuals. Of the collected risk assessments which modeled TRI chemicals, all used short-term modeling of chronic and/or acute effects, and provided quantitative risk estimates for the highest exposed individual rather than for the entire exposed population. Efforts continue to compare the Indicator model's relative rankings to those available for other groups of high priority chemicals. Development of the Environmental Justice Module of the Chronic Human Health Indicator may facilitate this type of analysis by addressing much smaller affected populations and geographic areas.*

It may be outside of your assignment from EPA, but it should be obvious that even the most rudimentary uncertainty or sensitivity analysis of the kind referred to above will have to include attention to validating the TRI emissions data. This would include attention to the accuracy of those emissions that are reported and to the compliance levels of the affected firms (in case under-reporting is a problem). Are you confident that EPA has these problems well in hand? (a)

*Response: The Indicators use the annually reported TRI data available to the public (from the "data freeze"). This self-reported data is validated by EPA. The Agency, which continues to place a high emphasis on its own data entry accuracy, is also engaged in outreach and training activities to provide assistance to facilities. The Agency has evaluated the accuracy of submissions from the outset of the program. Currently, EPA is conducting a data quality survey in cooperation with individual facilities representing a number of industry sectors. The Agency will use facility data to complete reporting forms for the 1995 data and compare its results with actual reports from those facilities.*

As the work group finishes this effort, I urge both EPA and the work group to devote some time to developing a defense to the "garbage in, garbage out" accusation. I don't think it is a fair charge but it is likely to be made given the large numbers of assumptions and guesses. One way to address it is to design some kind of reality check, which could be used to test the indicator before it is used as a gold standard. The reality check might be a paper study where a sampling of facilities are subjected to more detailed analysis or the check might involve some real monitoring to provide gross validation of some of the assumptions/predictions. If you move from one to several summary indicators, it is more likely that the "truth" will fall within the range of reported indicators. (c)

*Response: Such a "paper" validation of the Indicator method against site-specific risk assessments is planned. See response to first comment.*

**Agency (HQ) comments:**

We suggest that you rank some chemicals by the proposed method and compare the ranking to existing risk assessments for those chemicals to see if the proposed system distributes them in a similar order of concern. (d)

*Response: Such a validation exercise is planned.*

We are not presently well enough acquainted with the details of the methodology to offer advice on how to adjust it. The current ranking scheme of ordinal ranking lends itself to distortion at the high and low ends when the results are aggregated. A preferable weighing scheme should, if possible, correlate more closely with the actual risk. Perhaps the best way to choose a weighing scheme is to evaluate each option in a rigorous way so that the statistical properties are known. Sample runs or simulations, using various weighing schemes, would be helpful in evaluating each proposed weighing methodology. (h)

*Response: Ordinal ranking is not used in the method. Validation against actual risk assessments is planned.*

The calculated indices should be given a reality check by OPPT personnel prior to being adopted. (j)

This potentially very significant tool should receive notice and comment as well as peer review prior to its adoption. (j)

*Response: A formal review by the SAB and public participation in this process will address this point.*

**Agency (Regions) comments:**

In the May 29, 1992 Guidelines for Exposure Assessment, there is a brief discussion of using exposure assessment for Status and Trends (FR 22903). Accepted procedures and caveats for exposure assessments are also discussed in this guideline document. This document should be reviewed in the context of the TRI Environmental Indicator Methodology. (r)

*Response: In general, the principles applied in the Indicators method are consistent with the latest Exposure Assessment guidelines. However, there are recommendations in the guidelines (such as describing ranges of exposures, fully expressing uncertainty, etc.) that were considered too cumbersome for a risk-based, screening-level model.*

**Industry comments:**

It would be helpful if the document contained some discussion of the Agency's plans and next steps for review and evaluation of the draft report. Such a discussion would improve the understanding of where the report stands in the Agency's overall scheme. Such a discussion could also include a range of options available to the Agency for review, evaluation, pilot testing or demonstration of the proposed methodology with appropriate public comment. (s)

The document also suggest using the indicators as a prioritization tool. As such, the scheme would undoubtedly need to be validated so that relative rankings are meaningful (since scores are developed from mathematical manipulations of the data across numerous dimensions), each with their own set of uncertainties. Instead, the process could be tested using a small sample of the facilities and conducting a more detailed assessment, comparing it to the results as proposed using this methodology. (u)

*Response to above comments: This kind of uncertainty analysis will be part of the planned model validation task. Besides the formal review by the SAB, the Indicators model has been publicly demonstrated in a variety of settings (professional meetings, state organizations, Federal Facilities Roundtable, National Environmental Justice Advisory Council, etc.).*

We strongly encourage the Agency to provide additional opportunities for public involvement, including formally soliciting public review and comment, before this methodology is finalized. (v)

Review of this project by EPA's Science Advisory Board is highly recommended, as well as opportunities for additional public participation. (v)

*Response: Formal SAB review of the Indicators is planned. The 1992 public meeting and Federal Register notice solicited public comment, as reported here.*

### **Issue: Reporting and Interpretation of the TRI Environmental Indicators Results**

#### **Peer Reviewers comments:**

One "risk communication" issue-- I think EPA needs some guidance in how to investigate and explain changes in TRI indicators over time. Depending on the large national trends that will not be tracked as part of EPA's work on this particular project, changes in the TRI indicators could reflect real progress (or lack thereof) in reducing the volume of emissions (or in switching to less toxic pollutants), or they could reflect changes in national demography (people migrating towards or away from industrial facilities), changes in intranational or international competition (polluting firms relocating to other regions of the U.S. or to other countries), or the definitional changes in TRI referred to in Section VI of your report. In my opinion, EPA needs some guidance on how to investigate the causes of any "signal" it detects in the indicators, and in how to communicate to the public any mitigating factors that might explain changes in the indicators (or at least the qualitative uncertainties introduced by such factors). (a)

*Response: The structure of the revised Indicators model will allow for the diagnosis of causes associated with changes in the values. The Indicators allows calculation of sub-Indicators that allow tracking by chemical, industry, region, etc., and also allows separate model calculations (pounds only, pounds adjusted for toxicity only, pounds adjusted for toxicity and population only, as well as the full Indicator). These separate calculations allow the user to examine the causes of changes.*

What I urge that you focus on is the use of the indicators for communication. I believe that was what Administrator Reilly intended when he called for the development of such indicators. One of Mr. Reilly's first actions as Administrator was to commission *Reducing Risk*, and in a recent issue of the *EPA Journal* (March/April 1991; Vol. 17, No. 2) you will find him calling for national debate and discussion on using risk as a basis for setting environmental protection priorities. You should see your efforts in this context. (b)

*Response: The use of the Indicators as a relative risk-related priority setting tool is one of the major focuses of the work; this is discussed in the revised methodology documentation.*

I think it is a mistake to focus on one overall indicator for environmental risk. You mix apples, oranges, and all sorts of other fruits into a compote no one will find very satisfactory. It will be much better for you to focus on a "set of indicators," which is the term in the Statement of Work. Then for each one, you can discuss changes in a way that will be readily meaningful for the public.

Illustrative example:

The index of toxic air contaminants is down 25% for 1993 compared to the 1992 level, assuming the same evaluation of toxic substance emissions for the two years. However, of the 189 chemicals on the CAAA list, three more were found to be carcinogens as the result of recent animal testing by the National Toxicology Program, and two carcinogens had potencies reduced as the result of research on the biological mechanisms by which chemicals cause cancer in animals and humans. Changes in the evaluation methodology corresponding to these changes in the knowledge base for toxic air contaminants result in an even larger decrease in the indicator, from 25% to 28%. Using EPA's standard estimates for assessing cancer risks, the revised 1993 index indicates that up to 1440 annual cancer cases might be attributed to toxic air contaminants, compared to the 1992 index estimate of up to 2000 annual cancer cases.

Similar paragraph descriptions might be written about each of the other indicators in the set. Where an indicator corresponds to a program office in the Agency, the paragraph could provide a useful summary of how this program office is doing, by relating changes in the indicator to program activities affecting various source categories. The program office will clearly have strong interest in what this summary says. Where there is no program office, as for example, with indoor air, the paragraph description would be a good indication to Congress -- and the public -- of the importance (or lack thereof) of shifting more resources into this area.

A summary of which indicators moved up and which moved down would be instructive. Also, looking at the pattern of the changes could be useful, especially in comparison to where the Agency is placing its resources. The change in a weighted sum of all the indices might be least meaningful for strategic planning. (b)

*Response: The reviewer appears to be somewhat mistaken as to the purpose of this project. A major goal of the project is to use the TRI database (with its acknowledged limitations) to develop a set of Indicators as a tool to be used in support of relative risk-related priority setting. Separate Indicators for indoor air, mobile sources, etc., developed by other offices, could be used in conjunction with the TRI Environmental Indicators to set overall Agency priorities. This said, reports of Indicator results will always point out the relevant caveats. Reports could be written for the various sub-categories (chemicals, regions, industries) that may be examined with the current TRI Environmental Indicators. In order to accurately represent corrections in TRI reporting, reporting data for previous years is always re-calculated when the data for a new reporting year is added. It should be strongly noted that the Indicators do not estimate actual risk!*

**Agency (HQ) comments:**

It should be made clear that no rankings (conclusions) be available on a site-specific basis. (j)

*Response: OPPT believes that some site-specific analysis can be highly useful and meaningful, provided that appropriate caveats are addressed. It is, after all, the reporting of individual facilities which forms the basis of TRI aggregate numbers. Such analyses are to be viewed in the screening-level context of this tool and are not a substitute for site-specific risk assessment.*

Need to indicate how the reliability of indicated improvements will be defined. Would the reader (e.g., users, and policy makers) be able to readily discern that any declared "improvements" would not have addressed any of the risks posed to wildlife from exposure through their diets? If not, the policy maker could easily be misled. (l)

*Response: The limitations of the Indicators, and thus the interpretation of results, is discussed in the document.*

**Agency (Regions) comments:**

We suggest that a qualitative description of chemical release reduction (assuming we can be confident that TRI reductions are real) can provide an index of progress. The weight of evidence schemes could be presented with alphabetical designations, and reductions could be reported as numbers of pounds of releases toxicants in each category, or as reductions in the numbers of persons exposed:

- "Between 1997 and 1998, there was a 30,000 pound reduction in releases of class "A" neurotoxicants."
- "Four hundred persons exposed to class "B" developmental toxicants in 1987 are no longer exposed." (r)

*Response: The commenter's approach of categorizing by type of health effect does not utilize the greater amount of toxicity data which is available, nor does it take into account the relevant exposures to chemicals and size of the exposed population.*

**Industry comments:**

We in industry would welcome a clear and consistent system which provides a measurement of the health and environmental effects of emissions from our facilities. We hope that our personnel will be able to work closely with Agency staff in developing the EI's for our facilities. We suggest that companies would benefit by the release to the individual facility of its EI results at the chemicals and media levels. That would enable a facility to use the Environmental Indicators for itself much like the EPA will use them for a Region, to measure the

positive or negative change in its environmental impact. Another important benefit is that the facility would have the opportunity to validate the indicator numbers which have been calculated for it by the Agency. (v)

We understand that in addition to the composite numbers, disaggregated numbers will also be made available to the public. We predict that the public sector will be keenly interested in the Environmental Indicators as a potential tool to identify companies which do not meet their criteria of environmental stewardship. We think that the EI's which are calculated for each of the ten EPA Regions should be made public information, and we also would not be opposed to providing disaggregated numbers at the chemical and media levels on a regional basis. (v)

The TRI is being used today for purposes for which it was never originally intended. To the extent that these purposes are mandated and access granted to the public by law and regulation, we have no objection. The Environmental Indicators program, however, is being set up by the Agency for its own internal purposes of measuring environmental progress and establishing resource priorities. It has proposed to use the Toxics Release Inventory because it is said to be "the Agency's single best source of consistently reported emissions data." The choice of the TRI database for this purpose does not automatically carry with it the requirement to make all of the data generated under the program available to the public. In our view, for example, the fact that the TRI database is used does not confer on a community the right to know the calculated Indicator points for the facilities located there. We believe, therefore, that it would be inappropriate and unwise for the EPA to put facility-specific data into the public domain under this program, either directly or through freedom-of-information channels. It is unclear from the proposal whether individuals would have access to sufficient information to calculate Environmental Indicators on their own. Because of the huge potential for errors which might be introduced and propagated by self-calculators, we urge the Agency to take precautions to prevent this possibility. We further suggest that the project name has made an unfortunate link between the "Toxics Release Inventory" and the "Environmental Indicators Methodology" programs, a connection which is likely to create some false expectations. (v)

If the TRI indicators are necessary, they should be released to the public only as a national indicator. This would allow EPA to refer to one indicator rather than to many specific facility indicators for the measurement. National trends can then be established without specific facility location differences. Better techniques are available to measure an individual facility's impact on surrounding communities. (w)

In the SARA 313 releases and transfers report, a majority of the data is submitted as estimates. The report does not require additional monitoring or testing requirements. EPA's own estimates show only five percent of the facilities reporting use monitoring data to calculate emissions. Material purchased and hazardous waste data are the common sources of information for determining reporting requirements. Engineering estimates are used to determine releases in sequence batch operations within facility boundaries. Those estimates can vary between facilities because of available information and type of industry. The exposure potential of the chemical to the population can be inaccurate and the value of the comparisons between facilities will be low. Since the indicators measure the estimated impact of releases and transfers, it should only be considered a general measure of trends. EPA must recognize the limitations of the proposed indicators and develop indicators with less limitations. (w)

The resulting model will be so complex as to have little interpretability and relevance. For example, suppose the resulting index decreases by 20% from one year to the next for a particular region, state or company. Does this mean that emissions decreased 20%? Unfortunately, it does not. It does mean that existing predicted exposure indices decreased by 20% (based on some major assumptions), but this could have happened solely as a result of media transfer. The model appears to indicate that changing releases from air to water or to a landfill, or releasing it from a different location, (e.g., a different plant) can drastically change the index without changing actual emissions levels. Thus, changes in the index may have nothing to do with changes in emissions levels. However, this will not be clear without extensive analysis. In practice, what we are likely to see from one year to the next is a decrease in the index (assuming that TRI releases continue to decrease), without knowing what caused it. How much of the change was due to emission levels? disposal media? changes in product mix? location of emission sources? (u)

When companies, regions, or states are compared, it will be difficult to sort out the effect of the quantity of emissions from the effect of the exposed population size. If for example, the index is ten times as high in one state or another, is it because there are more emissions, because more populations lives near plants, or because the populations are the same size but some portion of the population lives closer (e.g., 100 m versus 1500 m) to existing facilities. This situation demonstrates the type of complex interpretive problems that will arise. (u)

*Response to above comments: The Indicators have been designed to reflect the relative risk-based impacts of TRI chemicals and the facilities releasing them, rather than assuming that all chemicals are equally toxic and that local populations are equally exposed to these emissions. This screening-level tool can help identify those chemicals, industry sectors, or facilities which may require further risk analysis. Performing analytical runs that can highlight trends in risk-related impacts and help identify the cause of such change is why the Indicators were developed. Such comparisons cannot adequately be made on the basis of pounds of emissions alone. The computer model is designed to allow for exactly these kinds of diagnostic runs. With this user-friendly pc-based, software program the user can readily analyze emissions to all media in terms of pounds alone, pounds with toxicity score, pounds with toxicity with population score, or the full indicator (toxicity-exposure-population). By looking at the differences in these results and disaggregating the data in various ways, the user can distinguish what elements are causing the changes in the indicator values year to year.*

### **Issue: Public Perception of the TRI Environmental Indicators**

#### **Peer Reviewer comments:**

The reviewer, by nature and training highly disposed toward quantitative methods, has had many negative experiences with the efforts by EPA and similar agencies to develop and use quantitative indices for problems that are inherently complex and uncertain... All too often these indices are misused, because managers ascribe a precision to them that is inappropriate (and frequently never intended by the developers). Rather than promoting understanding of the important issues of science that should be central to environmental risk management decisions, the quantitative indices often suppress these issues. Managers just want to look at the numbers. (proceeds to give extensive examples of misuse of cancer potency factors and the Hazard Ranking

System)... I am very concerned about potential misuse of the TRI Indicators. Your caveats and your description of the communication aspects of TRI need a great deal of expansion. (b)

*Response: Extensive caveats have been included in the methodology document and will be provided in both a User's Guide and in every analytical report using the Indicators.*

**Industry comments:**

One of the greatest concerns that comes to mind while reviewing the draft document is the potentially negative impact these indicator numbers may have on the general public. While it is emphasized that the environmental indicators are not intended to be quantitative, and are used as a tool to measure general trends related to the impacts of TRI chemicals, previous experience with TRI emission data has led us to believe that environmental groups will compile the data into a ranking system for industry, highlighting companies that supposedly pose the greatest threat to surrounding communities based on the greatest risk of contracting cancer, respiratory damage or other serious illness. Publication of these reports in a sensationalized manner in the news media could cause the public to become unduly alarmed and frightened, and in extreme circumstances, could even affect real estate values in certain localities. (t)

The environmental indicators document claims that its methodology is not a risk assessment method, and the methods do not calculate risk estimates in the absolute sense. Ultimately, the proposed indicator values will likely be used by local and state regulatory agencies and legislative bodies as true measures of actual risk. The public will perceive these indicator values in a similar fashion: as true risk measurements. (u)

While we believe EPA's intention to measure its own performance is commendable, and the technical quality of the methodology reflects considerable effort, we do not believe that the data will be used solely as a measure of the Agency's performance, nor do we believe the data will be straightforward enough to have any real interpretive meaning. (u)

We agree that the Indicators generated by the program should be a unitless value and reflect the general level of impact on a given area as a result of chemical releases from all of the sources within that area. However, there is some concern about how these Indicators may be used and potentially misused. Regardless of EPA's admonitions that these values are neither intended to be quantitative risk assessments nor a calculation of risk estimates, risk is nonetheless implied by their nature. We are certain that many in the public will view the results as a direct indication of risk and may become alarmed at the high numbers calculated. We fear that it will not be generally understood that significant error exists in the estimates due to the extremely conservative methods used to calculate the EI's. (v)

We therefore request that the Agency make every effort to be exceedingly clear in its instructions to the public on the meaning and significance of the Environmental Indicators. This might include a discussion of the kinds of data from which the elements are constructed, the substantial uncertainties associated with that data, lack of data in some cases, and the consequent conservative assumptions which are employed. Even more important, the EPA should specifically indicate which conclusions are valid to draw from the Environmental Indicator values, and which are not. (v)

We support development of meaningful data to assist in the public's understanding of SARA 313 reporting. Because the TRI indicators will become public, we are concerned about the potential misunderstanding and inaccuracy of the draft methodology. The public must understand that the indicators are not national environmental measurement tools, rather it is EPA's attempt to measure SARA 313 progress. (w)

The TRI indicator methodology is designed to interpret SARA 313 data. Because only certain manufacturing facilities are in the SARA 313 report, the indicators would not include mobile sources, agricultural releases, land uses, small to medium business, households, and SIC businesses currently not required to submit. Those sources can significantly impact human health and the environment along with the manufacturing facilities on the SARA 313 report. Since the public will view the indicators as a measurement of all sources, EPA should develop representative indicators that will not confuse the public or misrepresent data. (w)

*Response to above comments: OPPT believes the public is fully capable of interpreting risk-related indicators as well as reported results which currently address only pounds of emissions. All Indicator results are provided in terms of relative risk-based impacts. Because of the relative nature of this comparison of data elements, it is impossible to assign quantitative risk to an Indicator value (there is no fixed benchmark). This is a screening-level tool that can be used to identify situations with potentially high risk-related impacts that may deserve further scrutiny.*

### **Issue: Acute Effects Indicators**

#### **Peer Reviewers comments:**

Good to focus on four indicators, not one. Also, you are right to bound out spills and episodic releases. Make it clear that you are leaving out acute risks to health and the environment. To do this would require an enormous effort, given the nature of the time series data needed to do in order to do the job well. Don't expect this time series data from future reporting under the TRI system. (b)

*Response: The document makes clear that acute impacts are excluded at the present time because TRI reporting does not support such estimates.*

#### **Agency (HQ) comments:**

Focus of report is on chronic effects to health and environment at this time. Yet aquatic acute effects are included later on in the report. (l)

*Response: Acute effects data are only used as a surrogate for relative aquatic toxicity when chronic toxicity data are not available.*

#### **Industry comments:**

A major concern arising from the draft report is the potential for increased reporting requirements or redundant reporting. The amount of reporting required by EPA's various

programs has increased enormously over the last several years. This is evident in the TRI program as much as in many other Agency programs. In the last few years, there have been increased reporting requirements from the Pollution Prevention Act of 1990, as well as proposed expansion of the reporting list to include several hundred more chemicals. The benefit to the environment of this increased reporting is questionable. This additional information may be confusing, and even misleading, in many cases. While we support the attempt to make sense of the volume of data collected, we cannot support any further duplication, or additional reporting burden without substantial benefit to the environment. The Environmental Indicators document does make some attempt to make sense of existing information, but at the same time suggests substantial new reporting requirements. For example, the document recommends collection of peak release information. The Agency already has information, under CERCLA, on peak releases that occur above reportable quantities (RQs). If as stated in the document, the Agency wants to evaluate acute effects, it already has information for that purpose. We object strenuously to the collection of additional and redundant data. (s)

The draft discusses future consideration of developing acute environmental indicators, human health and ecological. One method that is suggested for determining acute effects would be to introduce peak release data into the TRI annual report. Considering the transitory duration of peak releases together with dispersion factors, it is difficult to conceive how any acute effect could be associated with any manufacturing operation. Indeed, if this were the case, such a health effect would have probably been detected in the facility or in the surrounding community. If a study of acute effects is desired, it would be more fruitful to compile information on health effects of catastrophic releases of chemicals, such as releases of chemicals in excess of RQs as defined by the CERCLA release reporting program. If acute health or environmental effects cannot be identified from catastrophic release of a chemical, peak releases would obviously have no effect. (t)

*Response to above comments: The current document only discusses the fact that acute data are not currently available, and mentions that they could be included in the indicator were they to be collected at a future date. It neither recommends nor discourages additional data collection, but simply describes the situation.*

## **Issue: Expand Methodology Beyond TRI Reporting**

### **Peer Reviewer comments:**

Don't just use the data in the data bases. Utilize expert judgment from the program offices inside the Agency and the scientific community outside the Agency to assure that what you are doing is sensible. In many cases you might pick up and use existing indicators and quantitative summaries - for example, the cancer risk estimates for air toxics done by the Air Office. If your numbers do not match their numbers and their judgments, you should find out why. Your computer implementation of the TRI indicators methodology should be sufficiently flexible so that methodology and data refinements can be readily accomplished. Trial use of your system and extensive interaction with various parts of the Agency (and the outside scientific community) should give you opportunities to make important improvements over the methodology and data you start with. I personally feel you should focus more effort on transformation, fate, and persistence in your methodology. Intermedia transfers need to be included, but I think you could waste a lot of effort trying to be comprehensive. The biggest concerns for health and environmental risk are usually the chemicals that are persistent (PCB's; mercury) or that transform into more toxic chemicals (chlorinated solvents into vinyl chloride through microbial action in aquifers; methylation of mercury). Very toxic chemicals that quickly degrade into harmless ones will be of concern only near the sources. Your TRI system ought to tell you that one of the biggest sources of human health risk from drinking surface water is the chlorination process, which produces trihalomethanes and lots of other chlorinated organic compounds, many of which are poorly understood and some of which are probably quite dangerous. Another important area with poor emissions data is the runoff of agricultural chemicals into surface waters and percolation into ground water. Indoor air is an underrated area for risks, because of the extensive use of toxic chemicals in home and office products/equipment and in building materials. The Agency focuses on radon and asbestos and largely ignores many other chemicals emitted in the indoor environment. Your system should indicate the need to shift priorities and resources in order to reduce human risk, even if some of the issues are outside EPA's statutory jurisdiction. (b)

You should find lots of similar insights from a careful reading of the main report and the supporting subcommittee reports for *Reducing Risk*. (b)

*Response: The reviewer makes good suggestions for the Agency as a whole; however, he misses the point that this is an Indicator based on the TRI database and is known to be limited by the inherent limitations in that data source (not all types of pollutants covered, not all sources of pollution covered). This Indicator should be viewed as one of several Indicators used by the Agency for risk communication, setting priorities and tracking environmental quality. The TRI Indicators Project has had a history of being responsive to suggested improvements that are in concert with its use as a screening-level tool. The Indicators have been designed as a flexible and adaptable concept.*

...site-specific risk assessment is hard and data-intensive. But by neglecting important aspects you can reach inappropriate conclusions. Do the risk assessment for a representative cross section of toxic releases (Statement of Work) at representative sites for releases into the appropriate media. Work with experts to find out how to do this properly. Then try to make the calculation simple and the results easy to communicate to the public. (b)

*Response: The purpose of the Indicators project is not to do risk assessment; in fact, there are objections to using the TRI data for risk assessment purposes, given the uncertainty in the underlying data. Furthermore, the issue of choosing "representative" chemicals was discussed at length by the Work Group and was rejected because of the difficulties of excluding particular chemicals or industries. The Indicators project is engaged in ongoing QA using site-specific risk assessment data to determine if the Indicators tracks risks in the same direction and degree.*

## **Issue: Selection of Chemicals and Industries**

### **Peer Reviewers comments:**

I note that since all but 1 of the 317 chemicals are either released in non-zero quantities (and may thus present individual risks that are worthy of consideration) or are not classified with low toxicity, that this argues for no exclusions at all at the outset. (a)

*Response: Currently no chemicals are excluded from the Indicator prior to determining if necessary data for modeling them is available.*

Do not try to maximize the comprehensiveness - adding as many chemicals and industries as possible. Rather, concentrate on doing a good job for a selection of the most important toxic substances of each risk class. The UJIA does not include all the industrial stocks on the NYSE, but rather a selected set of important ones. The Background Section in the Statement of Work talks about "capturing a representative cross section of toxic releases." That sounds like a good way to conceptualize what you are trying to do. My impression from the "Six Months Study" and subsequent attempts to assess (upper bound) national cancer risk from hazardous air pollutants is that most of the risk comes from a small subset of the chemicals - 20 to 25 of the 189 on the CAAA list. (b)

*Response: Due to the difficulties of selecting representative chemicals or industries, the approach of choosing a few "important" chemicals was not thought to be the best choice. Preliminary work with the Indicators has also demonstrated the risk-related impacts that a variety of chemicals and industries may have on the local level.*

Go for chemicals representative of those posing the highest risks. Don't limit yourselves only to the chemicals for which there is good release data. A valuable aspect of your system could be to go from observed concentrations posing high risk (in air, surface water, ground water, and soil), ask (the experts) where these concentrations might have come from, and then determine what additional data you need to get on emissions/releases in order to determine the source contributions more precisely.(b)

*Response: This comment is well beyond the scope and purpose of the Indicators project.*

Criteria 2,3,4, and 5 are simple screening rules, but for some chemicals they could lead to important omissions. Follow your idea under criterion 4 of circulating the list to get more data. Even small releases can lead to danger where there is high human exposure. Many proprietary chemicals will not be in IRIS, but OTS may have toxicological data on file under confidentiality

restrictions. Low toxicity compounds can give high risks with high doses. Do you know how many people are exposed to high doses of acetone? toluene? ethylene glycol? If you do not, find a chemist .. and ask him/her to explain what some of these chemicals in the Low-Tox list are used for. (b)

*Response: This comment refers to a section of the report where it was proposed to delete consideration of certain chemicals based on a given set of selection criteria. This section of the report has been deleted; all chemicals on the TRI roster are considered, although some may be excluded from calculation if toxicological and/or fate and transport data are unavailable to run the model.*

You might inquire why dichlorobromomethane and bromoform should definitely be in your system. (b)

*Response: All chemicals on the TRI roster are now included in the system; although, not all of these are fully modeled.*

Inclusion of facilities. I wonder if your decision to include all facilities makes sense. You will count lots of small ones, then ignore all those under the cutoff for reporting. Consider dry cleaners who emit perchloroethylene, or furniture makers who use paints and solvents. Will you get accurate estimates of total releases? (b)

*Response: It is not the intent of the TRI Environmental Indicators to capture all sources of pollution; rather it is to try to use the available TRI data in a risk-based way to help track trends and set priorities. Other indicators would be needed for other pollution sources.*

What is the justification for the specific cutoffs such as 1,000 pounds into each medium and 25,000 pounds of transfers? (c)

*Response: These cutoffs are no longer relevant, since they are no longer used as criteria to exclude chemicals from the Indicators.*

Exclusion of non-TSCA chemicals is not a very attractive option. Public is very concerned, for example, about pesticides. (c)

*Response: This is no longer a criterion for excluding chemicals from the Indicator.*

## Agency (HQ) comments:

Criterion 8 shouldn't exclude dichlorobromomethane. (d)

Under criterion 5, subcriterion b (oral and inhalation RfD greater than 0.1 mg/kg-day) is not a justifiable basis an RfD of this level may not be that high. The basis for the RfD needs to be included in the evaluation. Also, one needs to look at exposure levels. High RfDs could be exceeded by high exposures. (d)

Given some of the variability of reporting, you may want to reconsider dropping chemicals with no reports or zero releases in 1989. One alternative would be to drop only those chemicals with no reports or zero releases for all four years 1987 to 1990. The chemicals that were not reported in 1989 but were reported in some other year tend to have very small releases. But there are other chemicals with equally small volumes released that aren't being excluded from the Environmental Indicators (if they are of high toxicity), because they did report in 1989 (commentator attached table with reporting history of chemicals in Table 1 of the indicators report for all four reporting years). Using all four years instead of just 1989 wouldn't make that much difference one way or another - it might affect one or two chemicals. (i)

Commentator provided list of all TRI chemicals that did not have reports in at least one year, but have been reported on. Some chemicals were not reported early on, but were reported on in later years. This could reflect changes in chemical usage, or it could be due to errors in reporting. (i)

(Selection of Chemicals to Include in the Indicators) Excluding chemicals from this exercise seems like false economy; 87 + (unknown number of) chemicals may be excluded from TRI; the value gained by their exclusion should be clarified. Would very highly toxic compounds be excluded, even though firm understanding was lacking about their releases, and the nature of potentially exposed populations? (l)

Insufficient groundwork laid for proposing criteria for eliminating chemicals. (l)

Eliminating chemicals seems like false economy. I propose that OPPT evaluate all chemicals on TRI; this would result in a more balanced report. Chemicals with low toxicity will be apparent in final results. Policy makers can then use these obvious results. (l)

*Response to above comments: The current methodology does not explicitly exclude any chemicals from the indicator calculation, so there are no longer any "criteria" for exclusion. The intent is to include all chemicals on the roster if possible. However, it is still true that chemicals without reports are included in the indicator calculation as zeros. This is simply a practical matter of calculation: there is no attempt to distinguish true zeros from nonreports. Furthermore, resource and data constraints have not permitted toxicity scoring of all chemicals; therefore, certain lower priority chemicals are also excluded from full modeling in the Indicator calculation. These chemicals lack toxicity scores because Agency consensus toxicity data are not available for these chemicals, and resources were not available for compiling necessary data from the literature to derive scores. However, every TRI chemical is included in the Indicator*

*model on the basis of reported pounds of emissions, and trends and changes in reporting can be tracked.*

**Agency (Regions) comments:**

We do not support Criterion 1 as a viable criterion for selecting chemicals to exclude from the TRI indicator. We would prefer to work Criterion 1 as follows:

"Criterion 1: Exclude chemicals with no reporting."

We have found that small to medium-sized companies attending our seminars are still confused about what constitutes a release. In many cases, they have RCRA waste manifests on hand so they fill in the amount transferred offsite. However, they don't have similar data showing what amounts of the chemical were released to air, water, or land. Some of these people, especially if they don't have environmental training or adequate time to do the TRI reporting, will simply put down zeros instead of trying to estimate their releases. While we are attempting to address this problem through our workshops and data quality inspections, it's a slow process. We are only able to do about four or five data quality inspections a year in Region 8. (r)

For Criterion 2, the 25,000 pound offsite transfer trigger suggested in this draft is too high. We don't see a clear reason to have a different trigger amount for TRI chemicals transferred offsite. Just because those chemicals are sent offsite for proper treatment or disposal does not mean that they are properly treated and disposed. For example, a sewage treatment plant may not be able to treat some chemicals which it receives. In addition, Regional enforcement programs are kept quite busy dealing with improper disposal cases. (r)

The document contains a discussion of excluding "non-TSCA" chemicals from consideration. The term "non-TSCA" is never defined. Table 6 seems to indicate that the pesticides included in TRI are the "non-TSCA" chemicals. However, 16 of the 30 pesticides listed are on the TSCA inventory meaning that they have TSCA uses. In any case, whether or not a chemical is regulated under TSCA does not indicate its toxicity. TSCA status should not be used as a criterion for inclusion or exclusion. (o)

We support combining "relatively low reporting" with an evaluation of the chemical's toxicity and environmental fate data. If there are low releases and offsite transfers of a chemical, AND the chemical doesn't pose human health or environmental toxicity problems, then it would seem reasonable to exclude it from the TRI environmental indicator. Thus, we agree with the second criterion on page 15 if it is modified to delete the reference to transfers of less than 25,000 pounds. (r)

*Response to above comments: Same as previous.*

**Environmental Groups comments:**

The draft report suggests a number of different methods of choosing which chemicals to use as part of this set of "Indicators." Although EPA's resources might be saved by not ranking chemicals with no releases, this might lead to facilities switching to these chemicals and causing

releases in the future. Therefore spending the resources to rank all chemicals would be safer and produce a better guide for source reduction efforts. Chemicals should definitely be ranked if they have produced non-zero TRI reports, even if they only have non-zero releases of less than 1000 pounds. Chemicals should be ranked even if they are not part of the TSCA program, since others outside EPA would like to use this ranking scheme irrespective of EPA's statutory and bureaucratic structure. When future chemicals are added to TRI, they should also be added to the rankings. Since all components of the total ranking are chemical-specific, it will be simple to track "original" rankings consisting of the base year chemical set as well as the rankings containing all chemicals. (y)

*Response: Same as previous.*

### **Industry comments:**

Another proposal is that if no reported releases occurred for a listed chemical substance during the 1989 reporting year, it should be eliminated from consideration. This is not a valid method, if a chemical is of potential concern, it should not be eliminated simply because it was not reported to be released in any one specified reporting year. A scientifically based approach should be used in determining what chemicals should be eliminated from the chemical substance list. Also, it should be emphasized that before eliminating chemicals that currently do not have toxicological data, every effort to study and assess relative toxicity should be made. It is a risky assumption to exclude chemicals simply because it has been poorly studied. (t)

We support focusing the most attention on chemicals with the greatest potential for environmental impacts. By using a scientifically-sound approach to eliminate chemicals with relatively low toxicity concerns, the Agency can better focus its resources on the greatest potential impacts on the environment. There may be additional criteria (such as treatability, recyclability) that could be used to exclude chemicals from the indicator methodology. (s)

The concept of narrowing down the SARA 313 list of reportable chemical substances for the purpose of environmental indicators is advantageous, not only for the computational tractability mentioned in the draft report, but also for the benefit of concentrating on substances which potentially cause the greatest impact on both human health and the environment. (t)

The elimination of chemicals which have a toxicity weight below an established threshold, i.e., generally perceived to be non-toxic, is the preferred method for reducing the number of chemicals requiring impact analysis (Criterion 5, pg 13). It provides the opportunity for industry and government to focus attention to those chemicals with the greatest potential of harming the environment. (t)

Section III lists several methods of selecting chemicals to include in the calculation of the Indicators. It is sensible to limit the chemicals covered in this program to those which actually have significant impacts on human health, either because of their inherent toxicity or released quantity. To this end, we support Criterion 2 which excludes chemicals with relatively low (less than 1000 pounds to each medium) reporting histories, and Criterion 5 which excludes chemicals with relatively low toxicity concerns. In the case of Criterion 2, releases of this magnitude over the course of one year are extremely unlikely to have a major impact on health. Similar reasoning

leads us to exclude chemicals which have relatively low toxicity (Criterion 5), for example styrene, cumene, and toluene, since releases of these kinds of chemicals would also be expected to have little real impact on human health. Calculating Indicator values on these two groups of chemicals is not only unnecessary (and therefore wasteful of resources), but may suggest a higher level of concern than is appropriate. (v)

*Response to above comments: Same as previous.*

#### **IV. TOXICITY WEIGHTING**

##### **Issue: Toxicity Ranking**

##### **Peer Reviewers comments:**

The decision is made to equate exposure to carcinogens and noncarcinogens via a weighting scheme that equates, for example, a  $q_1^*$  value of 0.1 (the midpoint of the range 0.05-0.5) with an RfD of 0.001 (and  $q_1^* = 0.01$  equates to RfD = 0.01). The net effect of assuming that the two scales can be made proportional and of choosing the proportionality constant such that these equalities hold, is to equate the following risks: exposure at the RfD equals a cancer risk of  $10^{-4}$ ; exposure at 1/10 the RfD equals a cancer risk of  $10^{-5}$ ; exposure at 10 times the RfD equals a cancer risk of  $10^{-3}$ ; and so on. These are highly value-laden assumptions. The first of these three points of equality may make sense to the authors, but the 10,000:1 implicit weighting of the cancer:noncancer endpoints must at least be highlighted. And yet if this equality makes sense, it seems hard to justify some of the other equalities, either at the "high end" (exposure to 1000 times the RfD seems even worse than a cancer risk of  $10^{-1}$ ) or at the "low end" ( $10^{-6}$  risk may be de minimis, but according to some, exposure to 1/100 the RfD is trivial by definition since it is far below a threshold of effect). In addition to making much more prominent the implicit judgments buried in this system, you might consider maintaining two separate health indicators (one for cancer, one for toxic effects) so that there would be no need to linearize a probable threshold phenomenon, or use an alternative weighting scheme that equalizes risks at the "low end" only (see below). (a)

*Response: The relative scoring of carcinogens and noncarcinogens was taken from the Hazard Ranking System (HRS), the system used to rank Superfund sites. The Agency made a decision to equate carcinogens and noncarcinogens in this manner. The HRS rulemaking is a final rule and received significant public review during its development; therefore, while the relative rankings are subjective, they have been reviewed and made final under a different Agency rulemaking process.*

Assuming that noncancer chronic health risk varies with the ratio of the dose to the Reference Dose will be highly misleading when this ratio is less than unity. You should be finding out how many people may be exposed to levels near or above the RfD. (b)

*Response: It is true that if the ratio of exposure to RfD is much less than one, then risk is likely de minimis, and this will be reflected by a correspondingly low Indicator value. However, cumulative exposure to the same or multiple chemicals from one or several facilities, each with a*

*ratio of exposure less than one, could act additively to pose risk. Therefore, the choice has been not to drop them from the Indicator all together. The model is also not accurate enough to rely upon estimated exposure levels that fall just below the RfD.*

You have generally followed what others at EPA have done in constructing scoring systems. The results have some merit for crude ordinal ranking, but for some important chemicals the systems may give misleading results. (b)

I think there will be many problems with your scheme for both carcinogens and non-carcinogens. Many of these you may be inheriting from previous EPA systems like HRS. What you get is an extremely crude score that may useful some of the time. In other cases, look out. Try arsenic, a category A carcinogen that is one of the more common elements in the earth's crust. You will calculate very high potency scores for any water or soil in which there is detectable arsenic. (b)

*Response: Since the Indicator only calculates exposures associated with industrial releases, the contribution of background exposures is not relevant.*

#### **Agency (HQ) comments:**

A potential hazardous chemical is discounted due to weightings that make it seem less of a risk than it could be. (d)

*Response: The "discounting" of hazard is based on weight-of-evidence considerations that are consistent with EPA policy guidance.*

We recommend giving similar weights to ecological endpoints and to human developmental endpoints as are given to cancer. (e)

*Response to above comments: Weight of evidence is not needed with the ecological Indicator method because the toxic effects of chemicals on aquatic species can be directly observed. The weight of evidence for human developmental and other noncancer endpoints is considered during the RfD development process and therefore was not included again in the Indicators scoring system.*

#### **Environmental Groups comments:**

Community activists are often concerned with finding a better way to evaluate the relative importance of differing TRI releases. Many have complained that relying only on the number of pounds of all chemicals released is a poor way of prioritizing facilities for attention. Multiplying the numbers of pounds of different chemicals released and transferred by their toxicities would provide a better way of evaluating which facilities produce more pollution. The focus would not be on actual risk or exposure, which are difficult to quantify, but on a measure which would encourage pollution prevention and source reduction. The new "quantities of waste generated" numbers from section 8 of the new Form R could also be multiplied by their toxicities as a toxicity-weighted measure of waste production. Some community environmentalists might also find an EPA-created list of relative toxicity rankings useful for reference in various other

situations involving TRI data. To be useful in this manner, the toxicity rankings would have to be chemical-specific without being facility-dependent. The draft refers in one formula (pg. 47) to the toxicity ratings as being facility-dependent; hopefully this was a misprint or simply reflects that each toxicity rating will be multiplied by the pounds of chemical released or transferred. (y)

*Response: The equation in the current draft makes clear that toxicity scores are pathway-specific but not facility specific.*

Many community activists will disagree with even this set of rankings since they believe that this will encourage facilities to trade off one toxic chemical for another rather than striving to eliminate them altogether. In particular, it is important not to characterize the lower-ranking chemicals as "safe," particularly carcinogens. Under the proposed ranking scheme, a weak carcinogen would be given a low ranking compared to chemicals which caused other effects at low doses. Of course, EPA has stated that no level of exposure to a carcinogen is completely "safe" and that if possible these chemicals should be eliminated entirely. (y)

*Response: The Indicators use a reactive risk-related ranking, which is not based on some benchmark of toxicity. Also, the concept of threshold and nonthreshold effects existing for both carcinogens and noncarcinogens, is being evaluated within the scientific community.*

### **Issue: Basing Weight on Most Sensitive Endpoint**

#### **Peer Reviewers comments:**

"Applying such weights across categories of toxic endpoints would require a subjective evaluation of the relative severity of the health effects." (pg. 17 of draft) You reject the difficult subjective evaluation in favor of using the most severe endpoint - regardless of the relation of RfD to exposure. I predict a lot of your results will not make sense. Do some examples. How would you evaluate lead? (b)

*Response: An inherently toxic chemical will receive a high toxicity score based on the most severe endpoint. However, if exposure is low (relative to the RfD), then the corresponding Indicator score (which is a function of both exposure and toxicity) will reflect this fact.*

Is it really reasonable to assign the same weight to two chemicals when one has been shown to have several effects and the other only one effect? Your discussion was not very convincing on this point. (c)

*Response: The issue of weighting chemicals based on the most severe endpoint rather than on the number of effects was discussed at length by the Work Group. Because of the difficulty of assuring all effects were accounted for when scoring chemicals (especially poorly studied chemicals), the Work Group chose to focus on the critical effect instead. This is consistent with EPA's approach when developing RfDs for noncarcinogens.*

What risk level is used when a cancer endpoint is compared to noncancer endpoints? In what fraction of cases does cancer prove to be most sensitive endpoint and how sensitive is this to choice of risk level? (c)

*Response: The relative scoring of carcinogens and noncarcinogens in effect equates an exposure at the RfD to an exposure that yields  $1 \times 10^{-4}$  cancer risk. For the 1995 TRI reporting year (with a total of 606 TRI chemicals and chemical categories), 99 of the 340 chemicals and chemical categories with toxicity weights (29%) had a score based on cancer as the most sensitive endpoint in one or both exposure pathways. In some cases, the score based on non-cancer effects was equivalent to that based on cancer. Other chemicals may have had cancer effects that were not scored as highly as non-cancer effects, and would not be represented in the above count.*

**Agency (HQ) comments:**

Because of the numerous assumptions in the TRI methodology, we support a general health indicator with toxicity weights based on the most sensitive health endpoint. (f)

Additional information provided by sub-indicators for human health endpoints may not be worth the added complexity of analyzing separate indicators for each health endpoint. We recommend that the original indicator proceed as planned and that the results be analyzed to see if sub-indicators are warranted in the future. In the future, after the main indicator has been tested, the development of sub-indicators has the possibility of being useful for two reasons. First, the present indicator, by aggregating human health endpoints, may mask changes that the sub-indicators would reveal. Second, by reporting the health endpoints separately, the sub-indicator option would satisfy those who feel uncomfortable with the current method of ranking all health endpoints equally (i.e. carcinogenicity is ranked equal to chromosomal mutation). Reporting the health endpoints separately with these sub-indicators would result in contaminants with multiple endpoints being included in more than one sub-indicator. This "double-reporting" could in the future be a solution to the issue of how to weight contaminants with multiple endpoints. (h)

The lack of consistent data for multiple human health endpoints precludes any changes to the indicator at this time. In the future, when better data are available, one possible solution to the issue of weighing contaminants with multiple endpoints, as mentioned above, is to develop a sub-indicator for each endpoint, and report chemicals with multiple endpoints in each of the appropriate sub-indicators. This "double-reporting" would give extra weight to those chemicals if the sub-indicators are aggregated into one overarching indicator. Until better data on multiple human health endpoints become available, however, the issue of inconsistent weights for chemicals for multiple endpoints remains a problem. At this time, it is probably best to proceed with the current proposed weights until more data are available. (h)

**Agency (Regions) comments:**

The decision to apply toxicity weights based on only the most sensitive health endpoint limits the usefulness of the indicator. (m)

## **Environmental Groups comments:**

The draft ranking scheme would rank each chemical by the lowest level at which it causes any toxic effect, without adding to a chemical's ranking if it causes multiple effects. While this scheme does have the advantage of simplicity, one of the reasons given for it -- that many chemicals have not been thoroughly tested for a range of effects and that adding to rankings for multiple effects would lower the rankings of untested, possibly risky chemicals -- does not match the approach to uncertainty taken in the rest of the draft. As part of the toxicity ranking, ranks based on low weight-of-evidence effects are weighted down by a factor of 10. Chemicals for which there are no data are not ranked at all. This approach is consistent with ranking chemicals higher if they cause multiple effects, even if not all chemicals have been well tested for these effects. (y)

*Response to five previous comments: Only carcinogens in the WOE category of C have their toxicity weight adjusted downward by 10-fold. RfDs are based on a determination of the critical effect, with the idea that if one protects for that effect, than that value is also protective of any other effects occurring at greater exposures. The SAB will be specifically asked to provide comment regarding alternate strategies which could be used to account for multiple effects or to compare the severity of effects.*

## **Industry comments:**

Under "Selecting the Final Human Health Toxicity Weight for a Chemical," the document states, "The TRI Human Health Chronic Indicator method will weight a chemical based on the endpoint associated with the lowest dose at which an effect occurs." Although we believe we understand the intent of this statement, we would prefer to see it clarified by rephrasing to the effect that the Indicator should be based on the endpoint associated with the lowest dose at which an **adverse** effect occurs. This will prevent the Indicator from being based on an effect that has no significance to human health. (v)

*Response: The RfDs and RfCs are always intended to address adverse effects.*

## **Issue: Method Does Not Weight for Severity of Effect**

### **Peer Reviewers comments:**

The treatment of noncarcinogens is problematic because the approach does not account for the severity of the effect or how the severity (or frequency) of the effect changes with dose. More discussion of this limitation is appropriate. (c)

*Response: Additional discussion has been included in the methodology.*

The draft rejects weighting effects on the basis of their degree of severity as too "subjective." The draft should acknowledge that an arbitrary, if implicit, assignment of equal weights is also subjective. (Aside: Why not simply let different users set their own weights?) Lurking behind the work group's decision here might be an implicit bias in favor of hard science (e.g., emissions, toxicity, etc.) relative to soft science (e.g., psychology, economics, and utility

assessment). I did appreciate the references to the literature on health status indices and quality of life even though they were not incorporated into the indicator. (c)

*Response: The draft acknowledges that these decisions are arbitrary. While no suitable strategies for comparing severity of effects has been identified, the SAB has been asked to comment on this.*

### **Issue: Toxicity Data Used for Weighting**

#### **Peer Reviewers comments:**

Consider exploiting the statistical relationship between acute toxicity (LD<sub>50</sub>) and carcinogenic potency (TD<sub>50</sub>) to sidestep the problem of different endpoints and of potential carcinogens for which long-term bioassay data are not available. Although it is still controversial whether the observed correlation is artifactual or reflective of an underlying biologic mechanism (I can provide articles by Zeise, Crouch, Starr, Gold, etc., if you are interested), it may still be useful for decision-making and for providing incentives for data generation by the private sector. (a)

*Response: The current method adopts standard toxicological data sources upon which to base the Indicator.*

I would urge OTS to involve other agencies (especially NTP, NCTR, ATSDR) to ensure that all available toxicity data is assembled, not just that contained within EPA offices. (a)

*Response: To date, other agencies have not been involved in reviewing the supporting documentation for the Indicators toxicity data.*

EPA's cancer potency factors, as currently constructed, may be appropriate for screening purposes because it is believed that they provide an upper bound on the true but unknown carcinogenic effect at low levels of exposure. The same factors are not necessarily appropriate for inclusion in an index that purports to represent the status of environmental quality and risk. By combining a bounding estimate on cancer potency with other numbers (e.g., emissions) that are intended to be estimates of real-world quantities, the final scores are not interpretable. It is true that the proposed methodology is intended to capture relative risk and hence the absolute value of the summary indicator is not meaningful. If the cancer potency factor for each chemical were a similar bounding estimate (i.e, each possessed a similar degree conservatism or nonconservatism), then the approach would be valid for purposes of tracking relative risk. Unfortunately, this statement is highly questionable. One of the major themes of In Search of Safety: Chemicals and Cancer Risk is that EPA's cancer potency factors are far less conservative for some chemicals than others, even though they are calculated by a somewhat standardized procedure. I urge you to simply acknowledge this point in your discussion, since you have little choice but to use cancer potency factors. (c)

*Response: This point is discussed in the revised methodology documentation.*

...use default potency values for poorly tested chemicals; check with Alison Taylor at HSPH for a quick-and-dirty method based on acute toxicity tests and/or MTD. (c)

*Response: Instead of default potency values, OPPT chose to evaluate high priority chemicals and assign a weight based on available toxicity data.*

ED<sub>10</sub> and q<sub>1</sub>\* are highly correlated, and one can be essentially inferred from the other. (b)

*Response: Only q<sub>1</sub>\* is currently used in the method. The Benchmark Dose/Concentration approach to determining the effective dose is a fairly new concept. Of the current chemicals on IRIS, only five have assessments based on this benchmark approach. Four of these pertain to TRI chemicals or chemical categories. There is an RfD for methyl mercury, and RfCs for antimony trioxide, carbon disulfide and phosphoric acid. The IRIS values used in toxicity weighting for the methodology are based upon the most current IRIS data (April 1997).*

...acknowledge flaws in IRIS process; contact Kathy Rosica at CMA. (c)

*Response: The IRIS review process has undergone considerable change in the past several years. Generally, individual workgroups no longer conduct the reviews. Rather, as announced in the Federal Register several years ago, a pilot review of 11 chemicals was initiated; this review is ongoing. At that time public comment was solicited regarding this approach. As in the past, the public and industry may provide relevant information and toxicological studies to the review, but an IRIS submissions desk has also been established for these 11 reviews (as announced in the Federal Register notice). This submissions desk is maintained by the Risk Information Hotline in Cincinnati, Ohio (513/569-7254); the Hotline may be contacted for additional information. Each of these chemicals under review is assigned a manager and after preliminary review of data relevant to both oral and inhalation exposures related to cancer and non-cancer health effects, the review is sent through an Agency consensus process. In some cases, the Agency has elected to conduct this consensus review through workshops, and industry and the public have been directly involved.*

...how are the potency/toxicity of metabolites handled in the indicator? Am I right that you are focusing only on parent compounds? Shouldn't this be discussed more extensively (c)

*Response: The toxicity score is based on the compound listed in TRI, not its metabolites. However, to some extent, metabolites are considered when calculating the IRIS values and when the scores were assigned during the OPPT review process.*

**Agency (HQ) comments:**

(referring to idea of developing toxicity numbers where none exist) The cover memo (to HERD/CSRAD) uses the term "pseudo-RfDs. The word pseudo should be replaced with interim, provisional or unverified RfDs. Any unverified RfDs should be provided by OPPT staff analysis, though it is preferred that only verified values be used. (j)

*Response: The term estimated RfD is now used in the Toxicity Weighting Summary Document.*

EPA (not California)  $q_1^*$  values should be used. (j)

*Response: California values are no longer used as a basis for toxicity weighting.*

The proposed scoring systems for both cancer and non-cancer effects include both qualitative (weight-of-evidence or WOE) and quantitative elements (potency data). Although this approach is reasonable, several different methods are used as quantitative measures of potency ( $q_1^*$ , ED<sub>10</sub>, California  $q_1^*$  for cancer; RfD and RQ for chronic non-cancer effects), and qualitative WOE (EPA's Risk Assessment Guidelines and TSCA Chemical Scoring System), depending on the availability of toxicity evaluation. Since each evaluation method uses different criteria and assumptions, only Agency-wide assessment methodologies should be used for consistency. For chemicals that have not been formally evaluated by CRAVE or RfD/RfC work groups, interim or provisional assessments using the same procedures should be used until they are verified by Agency work groups. The methods described for the derivation of RfDs and cancer potencies (pages C4-C12) seem to follow Agency's procedures. (k)

*Response: All toxicity scores are now either based on Agency consensus values (IRIS or HEAST) or were derived from secondary data sources following the principles of Agency assessment methods. In the latter case, the toxicity scores derived were reviewed by an OPPT workgroup. Currently, this workgroup is evaluating those chemical with very high toxicity scores or those with large risk-related impacts in the Indicator.*

**Agency (Regions) comments:**

We suggest that the text on "no EPA toxicity data" be revised. We should more clearly explain that EPA may have toxicity information on these chemicals, but the Agency as a whole had not agreed upon an "approved" toxicity number. It is not accurate to say, as it does on page 13, that there are no "readily available toxicity data" on these chemicals. Because of the need to have Agency-wide agreement on toxicity numbers before information on a chemical is added to IRIS, there are many chemicals which do not have data in IRIS but which have been studied by the Agency over the years. In addition, some routes of exposure may not be covered by data in IRIS or the Heath Effects Assessment Summary Tables. For example, HEAST only covers oral and inhalation routes. It does not include dermal exposure. (r)

*Response: The current draft refers to the IRIS data as Agency consensus toxicity data. Dermal exposure is not addressed by the Indicators.*

We are confused by the text on Reportable Quantities (RQ's). This text seems to indicate that RQ's have not been developed for any of the chemicals in Table 4. However, methylenebis

(phenylisocyanate), CAS Number 101-68-8, and Dibenzofuran, CAS Number 132-64-9, each have an RQ of one pound because they were listed as hazardous air pollutants under Section 112 (b) of the Clean Air Act. You should have your contractor check the 1992 List of Lists to get the most current RQ's. If we have misinterpreted the meaning of Table 4, revised text should be written to clearly explain how IRIS, HEAST and RQ's are used to generate Table 4. (r)

*Response: RQs are no longer used as a basis of toxicity scoring in the current draft. RQs are based on a variety of effects and physical parameters. A much smaller number of TRI chemicals have RQs which are related to chronic health effects than those that have IRIS values. The range of scores for RQs is also narrower than that used for scoring in the HRS toxicity matrices, and the Indicators Work Group felt that the range of toxicity was more adequately described in the HRS approach.*

### **Industry comments:**

...under "Human Health Chronic Indicator" it is stated, "The current proposal is focused on general populations: individuals, particularly highly exposed individuals, are not the focus of the indicator. Additional indicators based upon highly exposed or sensitive subpopulations may be developed in the future." The methodology document proposes that the Reference Dose be used to determine the toxicity weight in calculating the element of the human health chronic indicator. It should be recognized, however, that the Reference Dose by definition is an estimate of an exposure to the entire human population, **inclusive of sensitive subgroups**. EPA's own definition is quoted on page 20 as "an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (**including sensitive subgroups**) that is likely to be without an appreciable risk of deleterious effects during a lifetime." This understanding is acknowledged in a later section of the document on page 25, where the application of uncertainty factors is justified in part by "extrapolating from studies on particular human populations to the general human population, **including sensitive subpopulations**." It is clear that the methodology has already thoroughly taken into account the effect of these chemicals on groups of greater sensitivity and therefore will not need to address this variable in the future. (v)

*Response: The footnote referred to the identification of exposed populations in the Indicator. While it is true that the RfD is protective of sensitive subpopulations, it may still be useful to have an Indicator that considers the size, location and/or level of exposure of such populations in conjunction with the toxicity score. Furthermore, some may find useful an Indicator that considers highly exposed, rather than sensitive, populations.*

The methodology document indicates that only EPA-reviewed scientific data, such as Reference Doses (RfDs) and Reportable Quantities (RQs), will be used when available. Although we appreciate that the intent of this is to ensure the use of high quality data, this policy will exclude a universe of valid, current scientific information. We believe that all available peer-reviewed scientific data should be utilized in the determination of human toxicity. (v)

*Response: The current draft includes toxicity scores that were based on review of readily available literature (from EPA health documents, ATSDR documents and HSDB on line information). However, review of the original literature for all chemicals is beyond the scope of this project.*

## **Issue: Weight of Evidence Discussion**

### **Peer Reviewers comments:**

The enclosed articles from the Journal of Policy Analysis and Management express my concern that the EPA alphabetical system of carcinogen classification may be misleading, and is probably inappropriate for use in quantitative potency adjustment. If Abt does not want to abandon such adjustments entirely, I recommend it at least compress its three categories into two (one for Groups A, B1, and B2, and one for Group C), with only a single factor of 10 separating A and C, rather than the current proposal for a 100-fold factor. This would not only reduce what I think may be a false dichotomy, but would make the indicator more stable over time, as chemicals shuttling between B1 and B2 status would not affect this alternative indicator. (a)

*Response: In response to this comment, the categories were compressed to two (one for A and B, one for C) with only a single factor of ten separating the categories.*

Weighting schemes should be approached with humility and great caution. Use of weight of evidence tables ... should be for crude ordinal ranking, not cardinal ranking. Simple severity times dose ... may be a good, starting point for non-carcinogens (subchronic and chronic, not acute). (b)

*Response: Many of the proposed weighting schemes in the original document have now been dropped; only the EPA Cancer Risk Assessment Guidelines weight of evidence scheme is still used. The purpose of continuing this approach is to remain consistent with current Agency policy on the evaluation of carcinogens.*

The draft does not adequately defend its treatment of uncertainty. Why is a factor of 10 appropriate for distinguishing known from possible carcinogens? For a different view of this issue, I recommend that you consult Adam Finkel of Resources for the Future. Why are the exposure uncertainty weights reasonable? (c)

*Response: The uncertainty factors chosen are arbitrary, as acknowledged in the document. the draft has been changed so that only a single order of magnitude separates the highest and lowest categories for both toxicity and exposure potential weighting. Additional planned model testing (for example, comparing test results to actual site-specific risk assessments) may suggest different values.*

Many risk assessors would question whether factor of 100 is appropriate for B1 vs. C. My guess is it should be less than 10. Might check with George Gray at HCRA, Adam Finkel at RFF, and Bob Moolenaar at Dow. (c)

*Response: A factor of 10 now separates A/B and C carcinogens.*

**Agency (HQ) comments:**

Category B1 should read: limited evidence from epidemiological studies and sufficient animal data. (d)

*Response: This correction was made.*

A category could be added for No evidence either way - similar to Group D. (d)

For the developmental effects numerical categories, a category "Not Studied" could be added. For mutagenicity, category H could have "or no studies conducted" added to the definition. (d)

*Response: These two comments are no longer applicable because explicit WOE schemes are no longer used in deriving toxicity scores for noncarcinogens. [See Appendix A of the methodology for a more complete discussion of this previous approach.]*

EPA Risk Assessment Guidelines should take precedence over OTS Scoring System if results for both are available. Guidelines are more recent and peer reviewed within the Agency. (j)

*Response: The weight of evidence classifications from the 1986 Cancer Risk Assessment Guidelines are now the only WOE classifications used in the document.*

TSCA Chemical Scoring System criteria for chronic/subchronic toxicity are proposed as a WOE tool. However, these criteria consider dose and severity of effect, not WOE. We've had the most problems with this set of criteria in the OTS Scoring System. These criteria need revision. Its use in its present form is not recommended. (j)

*Response: The TSCA Chemical Scoring System is no longer used.*

It is stated that for chemicals with RfDs, WOE is considered in the development of an RfD, and therefore, WOE is not used in this methodology for assigning toxicity weights for chemicals with RfDs (Executive Summary, page v). This is not necessarily true. A chemical can have a low RfD value due to large uncertainties resulting from lack of toxicity data (low confidence). On the other hand, another chemical can have a high RfD value even though it has been shown to cause adverse effects in several studies (high confidence). In our view, the proposed scoring system using the RfDs alone (page 36) does not adequately take into account the WOE. (k)

*Response: The comment is referring to the fact that in developing an RfD, low confidence in data can result in a lower RfD because more uncertainty and modifying factors are often used to adjust insufficient data. A lower RfD would give a higher toxicity score under the Indicator method. This seems counter to the Indicator approach, which gives a lower score to carcinogens with less strong weight of evidence. However, weight of evidence is considered differently in the development of noncarcinogens. The question of whether a chemical causes a noncancer health effect in humans is considered qualitatively in the selection of the critical effect that is the basis*

*for the RfD. The selection of critical effect is itself a weight of evidence determination. The uncertainty factors and modifying factor are then used in the potency determination. That is, UFs and the MF not used to adjust for weight of evidence that the effect is relevant in humans (as cancer WOE categories are), but rather to estimate conservatively the dose above which humans may be adversely affected.*

It should be noted that EPA's WOE Category for carcinogens and non-carcinogens is not "equivalent" to TSCA Chemical Scoring System. Please see the listed criteria on pages 23-27. As discussed above, TSCA Chemical Scoring System should not be used at all. (k)

*Response: The TSCA Chemical Scoring System is no longer used.*

**Agency (Regions) comments:**

The authors should note that weight of evidence schemes can be designed to provide evidence that a chemical causes a specific health effect in general or specifically in humans. The categories for carcinogens are related to likelihood of effects in humans. (r)

*Response: This point is now noted in the document.*

**Industry comments:**

The toxicity weighting matrix for potential human carcinogens in Figure 4 (page 34) considers only two categories: A/B (known/probable) and C (possible). We feel strongly that it is more correct to divide all potential carcinogens into three groups, not two. Under the proposed TRI EI Methodology, a known human carcinogen (Category A) would be weighted the same as an animal carcinogen (Categories B1 and B2). To do so runs counter to the mainstream of scientific thought and is inconsistent with the use of carcinogen classification by other EPA offices. The carcinogenicity of Category A chemicals is based on human data and consequently should be weighted more heavily than that of Category B chemicals, which is based primarily, if not entirely, on animal data. (v)

*Response: The recommendation to combine A and B carcinogens was a recommendation of two of the original scientific peer reviewers of the TRI Environmental Indicators method. Furthermore, the 1986 Cancer Risk Assessment Guidelines state that for regulatory purposes, A and B carcinogens have adequate weight of evidence for regulation, while C carcinogens generally do not. Therefore, A and B categories have been left combined in the methodology.*

## **Issue: Other Related Comments**

### **Peer Reviewer comments:**

What are these chemicals used for? [Table No-Rep] Is there a potential for human exposure except through direct ingestion of a food or drug? (Not likely for the food coloring agents.) Ask the experts or look up the chemical in a handbook. DBCP is an example of a pesticide no longer in use that caused very serious groundwater contamination, which still persists. (b)

### **Agency (HQ) comments:**

Potency values are route-specific for both  $q_1^*$  and RfDs. The approach violates current practices for use of  $q_1^*$ s, RfDs, etc. Extrapolation of toxicity parameters are done without consideration of route of exposure. (j)

*Response: Where route-specific values are available, these are used. However, where data are unavailable for one exposure route, and in the absence of data indicating differences in route-specific toxicity, the toxicity score from the other pathway is used. Although route-to-route extrapolation is not current EPA policy for establishing health reference values, it is occasionally performed in the absence of key data and with appropriate precautions (although it may be performed with appropriate precautions when key data are missing), it was deemed appropriate for a screening exercise such as the TRI Environmental Indicators project that uses only order-of-magnitude toxicity scores. This approach is consistent with the HRS methodology for developing toxicity factors.*

... describes approaches to WOE for genotoxic and developmental effects (Our comment is not on the WOE schemes but on the more fundamental question of whether these effects belong in a chronic toxicity indicator): these risks cannot be evaluated adequately if only long-term average exposure estimates are available. Like acute effects, information is needed on short-term exposures. Additionally, most genotoxic effects lack meaningful dose-response data and are also considered to be non-threshold effects. (j)

*Response: Even though short-term exposure is the type of exposure of concern for these effects, exposures do not have to be at extremely high levels to pose a concern, since many developmental toxins have short-term risk values that are in the same range as chronic systemic toxin risk values. Since this is so, using the long-term annual average exposures can be used roughly to identify potential areas of concern, if one assumes that exposure levels vary over time and that there will be peak exposures that are as high or higher than the long-term average.*

*Genotoxicity is considered a non-threshold effect under most circumstances (germ cell versus somatic cell nucleic acid damage). Consequently, a time-weighted average exposure may provide a reasonable indicator of toxicity. It is recognized that at high exposure levels DNA repair mechanisms may not function appropriately. However, DNA damage is known to occur at low exposure levels. In addition, cell death often occurs with high exposures, preventing the perpetuation of damaged DNA in the cell line. For purposes of obtaining indicators of genetic toxicity, it is assumed that low level exposures may cause genetic damage. The damage which*

*may occur during peak exposure periods is of interest, but it is not possible to evaluate such risks with the available exposure data.*

Regarding the development of risk scores for cancer vs non-cancer effects: More consideration should be given to the presumed threshold for non-cancer effects. It may be more appropriate to develop scores for non-cancer risks by looking at the ratio of the exposure level to the RfD (or RfC) rather than multiplying individual scores for the two factors. Any exposure below the RfD is essentially a de minimis risk and should receive a score comparable to a de minimis cancer risk. The real issue with noncancer effects is how to estimate risk at exposures between the RfD and known effect levels. EPA has not been able to do this as yet. (j)

*Response: Since the toxicity score of a chemical is inversely proportional to its RfD, multiplying its toxicity score times the exposure score will yield a value that is similar to the ratio of exposure to RfD (that is the value will change in the same direction and degree as the exposure/RfD ratio). The method does not assign exposures below the RfD a de minimis risk level, but assigns lower Indicator values to such exposures. This is because even if exposure to any one chemical is below a toxicity threshold, there may be simultaneous exposures to multiple chemicals that together could pose a risk, or multiple exposures to the same chemical emitted from several nearby facilities could occur. Due to the lack of site-specific data, modeling of exposures is not considered to be accurate enough to differentiate between exposures near the RfD. Finally, it is true that for most chemicals the risk shape of the dose-response curve between RfD and known effect levels is not known; as stated in the draft, it is assumed that risk varies as the ratio of the estimated dose to the RfD.*

## **V. TRI CHRONIC ECOLOGICAL INDICATOR**

### **Issue: Indicators Do Not Address Terrestrial or Non-Water Column Wild Populations**

#### **Peer Reviewers comments:**

Why does the draft focus exclusively on aquatic toxicology? (c)

*Response: As discussed in the text, there are inadequate data to characterize exposure to terrestrial populations; however, as a rough surrogate, the human chronic health Indicator could be used as an Indicator of potential risk to terrestrial mammalian species.*

#### **Agency (HQ) comments:**

Need to make clear to readers and users that the method does not address wild populations. The proposed "indicators" scheme is designed for use in evaluating successes in reducing environmental risks. Meaningful declarations of success must be defined. However the method does not address impacts on wild populations because of the acknowledged complexity of that task. For this reason, any "successes in reducing environmental risks" should be clearly defined in terms that unmistakably indicate that wild populations were not addressed. (l)

*Response: Caveats regarding the limitations of the Indicators are discussed in the text.*

Chronic eco impacts will address only the aquatic water column. Need to address sediments and terrestrial environments to provide meaningful indications of success in reducing environmental risk. (l)

With respect to the Ecological Chronic Indicator, you need to address terrestrial plants and animals in order to be credible indications about the environment. This is critical. If you purport to develop a scheme for evaluating successes in reducing risk, you need to know about the size of populations being exposed. (l)

With respect to Toxicity Weights -- Ecological, you should use ecological data to address impact on terrestrial plants and animals. Human health data will not pick up most effects to plants; and effects from some chemicals may not be significant for humans, but it will be a major problem for wild species. (l)

*Response: While these suggestions would add to the usefulness of the Indicators, these efforts are beyond the current scope of the project due to their complexity, data gaps, and available resources.*

### **Issue: Ecological Impacts**

#### **Industry comments:**

The scientific quality of this section prompts similar concerns about the data and the way in which it will be used to develop an EI for chronic ecological effects. We feel that this section needs significantly more development to approach a rigorous description of chemical impacts on ecological systems. For example, the authors suggest that the Environmental Indicator values developed by the proposed methodology accurately reflect ecological impact. However, there is no discussion and no measurement of ecological impacts associated with specific EI values. Discussion of environmental receptors, i.e. organisms in the environment, has been neglected. Without this element, the EI values can only reflect estimated exposures and potential impacts. The critical link - associating EI values with real environmental impacts - is thus omitted. Furthermore, there is no validation that the proposed methodology can even quantify real exposures with any accuracy. We believe that the methodology needs to be improved if the EI values are to be credible. (v)

*Response: As with the Chronic Human Health Indicator, any Ecological Indicator developed would be intended only to be an indicator of potential impacts, not an assessment of actual risk. This is stated several times in the document.*

## **Issue: Ecological Toxicity Data**

### **Agency (HQ) comments:**

Water quality criteria are guidance levels and carry no regulatory weight. The same for aquatic RQ's. (1)

*Response: It is true that water quality criteria are guidance levels; however, this is considered to be adequate for the purposes of the Indicator method.*

(Toxicity Weights -- Ecological) Do QSAR calculations where data are lacking. (1)

*Response: This is an option when the Ecological Indicator is developed.*

(Toxicity Weights -- Ecological) Use QSAR calculations where few data are available on terrestrial or avian species. (1)

*Response: This is an option if a terrestrial Ecological Indicator is developed.*

The use of Acute or Chronic Ambient Water Quality Criteria (AWQC) or other criteria developed by the Office of Water may be inappropriate for use as criteria in this proposed ranking system. See comment for p vi. Likewise, this document specifies in the Executive Summary that only Chronic values are to be considered at this time; therefore, the use of Acute data from Water Quality Criteria is out of place. If a preference is to be set, then measured data from existing databases, such as AQUIRE, should be used, followed by estimated toxicity using Structure Activity Relationships and others should be left blank pending development of data. (1)

*Response: Although on first inspection it may appear to make more sense to rank annual TRI chemical releases using some indicator of long-term (chronic) effects on aquatic organisms (e.g., chronic AWQC or other chronic toxicity data), in actuality, it is not known whether TRI releases are truly chronic or acute in duration. Preference was given to the use of chronic toxicity endpoints simply because chronic effects (measures of growth, reproduction) are generally more sensitive than acute effects (usually measures of lethality), and thus provide the most effective screen for aquatic hazard. However, chronic toxicity data of appropriate quality and exposure durations are far less abundant than acute toxicity data. Therefore, a much smaller number of TRI chemicals could be ranked using chronic toxicity data than acute toxicity data.*

*In the absence of acceptable chronic toxicity data for a particular set of chemicals, the procedure is to derive relative hazard rankings based on comparative measures of acute toxicity (acute AWQC or other acute toxicity data). This is consistent with the established goal of the TRI indicators, which is to provide a consistent basis for ranking hazards for a given set of chemicals. While it is true that relative rankings of acute and chronic sensitivity may differ somewhat across chemicals (i.e., the relationship between acute and chronic toxicity varies across species and chemicals), this does not invalidate the use of acute toxicity comparisons as a method for determining relative hazard rankings for a particular group of chemicals. Furthermore, recent models have been developed that quantitatively relate acute and chronic*

*toxicity responses of aquatic organisms (Barnthouse et al., 1990; Mayer et al., 1994; Lee et al., 1995). These studies suggest that reliance on acute toxicity data (in the absence of chronic toxicity data) for establishing relative hazard rankings is not unreasonable or toxicologically unsound.*

*References:*

*Barnthouse, L.W., G.W. Suter II, and A.E. Rosen. 1990. Risks of toxic contaminants to exploited fish populations: Influence of life history, data uncertainty, and exploitation intensity. Environ. Toxicol. Chem. 9:297-311.*

*Mayer, F.L. G.F. Krause, D.R. Buckler, M.R. Ellersieck, and G. Lee. 1994. Predicting chronic lethality of chemicals to fishes from acute toxicity test data: concepts and linear regression analysis. Environ. Toxicol. Chem. 13: 671-678.*

*Lee, G., M.R. Ellersieck, , F.L. Mayer, and G.F. Krause. 1995. Predicting chronic lethality of chemicals to fishes from acute toxicity test data: Multifactor probit analysis. Environ. Toxicol. Chem. 14: 345-349.*

**Industry comments:**

The methodology document seems to assume that a large body of data is obtainable on the TRI-listed chemicals, data which we believe may not exist. The authors have suggested that many values of aquatic toxicity and bioconcentration are available, without appearing to have evaluated either their accessibility or merit. For example, some TRI chemicals have published data points which vary by 1000-fold. The EI methodology needs to provide guidance on how data will be selected for calculation of the EI's, and we reemphasize that all peer-reviewed data should be admissible to the process. (v)

*Response: For some chemicals the Office of Water has evaluated and selected toxicity data for use in its programs. For these chemicals, Office of Water data can be used. For other chemicals, extensive collection, review and evaluation would be required.*

**Issue: Ecological Toxicity Weighting Approach**

**Peer Reviewer comments:**

Given my (rudimentary) understanding of ecology, I would give big weights to persistence and bioaccumulation. As a consumer of your system, how do I know whether your assumptions are reasonable? Also, shouldn't the emission weights reflect the baseline health of the particular ecosystem that is about to receive the emission? Some of the chemicals excluded for low toxicity to humans may have significant aquatic toxicity (e.g., copper and zinc). Please double-check this matter. Should you mention non-aquatic organisms or why they are excluded? (c)

*Response: Bioaccumulation is considered in the aquatic toxicity score. These weights are based on the HRS system, a reviewed and published EPA rulemaking. Emission weights cannot reflect*

*the baseline condition of the ecosystems, since these data are not generally available even for site-specific risk assessments, and are definitely not available on a national level for the purposes of the Indicator. When the aquatic Indicator is developed, the entire roster of TRI chemicals will be considered for aquatic toxicity; the chemicals included in the Ecological Indicator may differ from those included in the Chronic Human Health Indicator. As discussed in the text, there are inadequate data to characterize exposure to terrestrial populations, and these populations are excluded.*

**Agency (HQ) comments:**

Inherent toxicity (hazard) does not change with time. Exposure or reduction in the amount of release/unit/time can reduce the water concentration, thus reducing the probability of the hazard level being realized in the aquatic system. The same effect could be accomplished by determining the inherent toxicity with actual test data or in the absence of test data, use Structure Activity Relationships (SARs) to provide estimates of toxicity followed by ranking according to hazard levels. (1)

*Response: The Indicators Work Group disagrees with this comment. Both hazard and exposure (including chemical fate and transport) are considered important in the scoring process, so both remain in the suggested approach.*

[Toxicity and bioconcentration] should be considered separately because toxicity and bioaccumulation have unequal weight and importance. (1)

*Response: Separate values are developed for bioaccumulation and toxicity, and are only combined after the chemical is scored separately on each factor.*

Extreme caution should be exercised when attempting to combine weights of toxicity and bioconcentration. High bioconcentration is of little consequence if the toxicity (hazard) is nil to low. Toxicity of compounds through bioconcentration is important via indirect effects on other organisms in the food chain. (1)

*Response: A high bioaccumulative but low toxicity chemical would yield a relatively low score which would reflect its lesser consequence.*

**The following comments refer to aquatic toxicity weighting matrix:**

[There are] missing values in water solubility column. (1)

*Response: These values are not missing. The values assigned follow the Hazard Ranking System scoring for bioaccumulation potential. In this system, the score for chemicals with water solubility greater than 1500 mg/l is three orders of magnitude lower than for chemicals with solubility of greater than 500 to 1500 mg/l.*

Under weight why not use 0.1, 1, 10, 100 etc instead of 5's? (1)

*Response: These values follow the HRS scoring system. However, the actual value of the score is irrelevant as long as the relative score is maintained.*

How are water sol. values in the first three related? Or have they been correlated by looking at various chemicals as examples? (1)

*Response: The specific grouping of categories and assignment of scores was done as part of the HRS process and is documented as part of that rulemaking (see 55 FR 51532 and docket). In general, however, the measures chosen and scoring assigned were based on an evaluation of scientific data relating bioaccumulation potential to those particular chemical properties.*

Why does log Kow stop at 6.0? Do chemicals with log Kow values >6.0 pose no risk? (1)

*Response: For chemicals with log Kow > 6.0, the score should be based on the water solubility.*

All chemicals with water solubility <25 mg/L do not necessarily pose a high bioconcentration potential. (1)

*Response: This statement may be true, but the water solubility value would only be used as an indicator of bioaccumulation potential if more specific data (BCF, log Kow) were not available.*

The Aquatic Toxicity Weighting Matrix. Again using the aquatic toxicity value and the bioaccumulation potential value to determine the aquatic toxicity weight is not recommended. As mentioned previously, just because a chemical has a high log  $K_{ow}$  does not mean it is toxic. (1)

In the first Table no justification is provided for developing any of the six criteria levels shown in the Table. In actual practice in OPPT the ranges used for scoring log  $K_{ow}$  with respect to bioconcentration potential are (1):

$K_{ow}$	<3.5 is low concern
$K_{ow}$	3.5 to 4.3 is of moderate concern, and
$K_{ow}$	>4.3 is of high concern.

*Response: When the ecological indicator is fully developed, the issue of an appropriate de minimis can be resolved. However, it should be noted that a relatively low toxicity chemical may be of concern if it bioconcentrates to high enough levels.*

Aquatic Toxicity Weights. First, the Executive Summary states that only Aquatic Chronic effects will be used in the scoring system at this time. Second, what is the source of the "Life Cycle Chronic NOAEL Criteria". Third, the "Chronic AWQC or AALAC" are for guidance only from OW and may not be appropriate for use by OPPT. Fourth, the standard practice is to report aquatic toxicity values in mg/L and not ug/L. Fifth, the normal practice for ranking aquatic hazards in OPPT is well established according to the following criteria:

< 1 mg/l - High  
1 to 100 mg/l - Moderate  
> 100 mg/L - Low

For the scoring of Chronic toxicity values, the above values are divided by 10. (l)

*Response: First comment: As stated earlier, acute data would be used in the absence of chronic data to approximate relative toxicity. Second comment: life cycle or chronic NOAELS would be derived from the literature through a search of AQUIRE. Such data would only be used after a review by OPPT staff. Third comment: AALACs have been dropped. Use of guidance from OW was considered appropriate for this screening-level exercise. Fourth comment: The use of µg/l is consistent with the units used in the HRS. Fifth comment: the OPPT ranking system is qualitative in nature and thus is not compatible with the current method.*

(Aquatic toxicity table of weights) Clarify whether acute LC50s will have the same weight as LC50s derived from prolonged testing, e.g., 30-day LC50s. (l)

...by using the product one obtains large numbers that can result in error if zeros are dropped by accident. A matrix value of 500,000,000 may be so large as to be meaningless to a policy person. (l)

*Response: These details can be resolved when the ecological indicator is fully developed.*

#### **Industry comments:**

A great deal of weight is given in the methodology to evaluating aquatic toxicity based on bioconcentration factors. However, the proposed approach utilized older (regression) models to determine toxicity factors; we believe that more current techniques would enhance the accuracy of the values. Furthermore, the bioconcentration weighting scheme is constructed in such a way that a substantial contribution to the indicator value will be made by **any** water-insoluble substance, **whether or not it has been shown to be toxic**. We are sure that this was not the intent of the authors, since the indicators are intended to provide an index of negative impact on the ecological community. (v)

*Response: Other methods could be employed to assess toxicity when the ecological indicator is implemented. Furthermore, it is true that a chemical will contribute to the indicator if there is high bioconcentration but low toxicity. However, chemicals with no evidence of aquatic toxicity whatsoever could certainly be excluded for the indicator. The issue of setting a de minimis toxicity level could be further explored when the ecological indicator is fully developed.*

## **Issue: Calculation of Bioconcentration Factors**

### **Agency (HQ) comments:**

Very little test data are available on measured BCFs. However, the BCF SAR of Veigh and Kosian can be used to estimate the potential of a chemical to bioconcentrate based on the use of Log K<sub>ow</sub>. It must be remembered that this provides an estimate of the "potential" bioconcentration and as such should be used as a screening tool. Such factors as metabolism, biodegradation, and non-toxicity of compounds with high K<sub>ow</sub> values affect the ultimate bioconcentration factor. (1)

The AQUIRE database has limited data on measured Bioconcentration Factors. Since Log K<sub>ow</sub> can be estimated for most discrete organics using the computer program CLOGP vers. 3.3, the BCF SAR developed by Veigh and Kosian can be used to predict the potential BCF for most organic compounds. (1)

Are we all agreed on the model which will be used to develop bioaccumulation factors (not bioconcentration factors) for aquatic organisms? (1)

*Response: These details regarding the calculation of BCF can be resolved when the ecological indicator is fully developed.*

## **Issue: Weight of Evidence Considerations**

### **Agency (HQ) comments:**

(Toxicity Weights - Ecological) Weight of evidence (WOE) may be needed to address the problem of uncertainty about which toxicity values to use. There may be no problem where AWQC are available; but where they are lacking, e.g., in using AQUIRE, you probably will need to review individual studies to arrive at a reliable value. Where the available studies are not reliable, WOE may be necessary. (1)

*Response: If data from AQUIRE is needed to implement the ecological indicator, a review of data by OPPT experts would be required. This review would incorporate considerations of the quality of the data base. (This is different from typical WOE evaluations, which assesses whether the impacts observed in the available toxicity literature have relevance for the species of concern in the assessment, e.g., humans).*

(Toxicity weights - Ecological) Eco does not use WOE and this discussion of this term should be excluded from the discussion. (1)

*Response: The text now indicates that WOE is not considered in the ecological indicator method, because aquatic toxicity data provide direct evidence of the presence or absence of an impact on aquatic species.*

## **Issue: Ecological Exposure Weighting**

### **Agency (HQ) comments:**

(Exposure Weights - Ecological) Using ambient water concentrations of the chemical to define potential aquatic exposure poses a problem; it ignores exposure from the diet, and also from sediment, especially for chemicals with low water solubility and high Kow. (l)

*Response: It is true that using water concentrations do not capture bioaccumulation concerns. However, the bioaccumulation is considered as part of the toxicity weight given a chemical.*

Clarify meaning of "equally vulnerable locations." Does it mean that all locations are assumed to contain large numbers of endangered species? Or does it mean that all locations are assumed to contain only small numbers of species, most of which are judged to be pests (e.g., those often found in streams receiving industrial effluents)? (l)

*Response: Since the data are not available to distinguish between locations with vulnerable characteristics (such as presence of endangered species) and those without, the method de facto assumes that chemicals are released to "equally vulnerable locations". This fact was mentioned in order to point out an uncertainty in the analysis.*

## **Issue: Calculation of Ecological Indicator**

### **Agency (HQ) comments:**

Why not the "sum" or average of these weights? (l)

*Response: The individual component (toxicity, exposure) scores are multiplied to yield an indicator element because it is assumed that toxicity and exposure act in a multiplicative way (i.e., in the risk assessment paradigm, risk = toxicity \* exposure). However, indicator elements are then summed, since risks from various locations and from different chemicals are assumed to be additive.*

### **Industry comments:**

Finally, the methodology assumes that the individual Indicator values calculated for each chemical should be combined by summation to determine the final cumulative Environmental Indicator of chronic ecological effects. This procedure runs counter to the EPA's own NPDES technical support document, which explicitly rejects the assumption that chronic effects are additive. We suggest that the scientific rationale for method be reviewed to discover whether a more valid means of combining results can be found. (v)

*Response: In addressing this issue, it should be brought to attention that the intent of the ecological indicator is to provide relative indications of ecological hazard and not to provide absolute measures of ecological risk. Thus, in order to provide relative hazard rankings of different groups of chemicals, the individual ecological indicators were summed across*

chemicals. The summation of individual hazard rankings assumes that the individual hazard rankings of chemicals is additive. This assumption follows the logic that greater hazard would be associated with a group of, say, ten chemicals each of equal toxicity to aquatic organisms compared to hazard associated with a single chemical from the group.

From reviews of the literature on the acute toxicity of chemical mixtures, the assumption of additivity is generally supported (see U.S. EPA, 1991, p. 24; Suter, 1993, p 234-238). However, there appears to be far less information available on the additivity of chronic toxicity of chemicals to aquatic organisms. As reviewed by U.S. EPA (1991), Alabaster and Lloyd (1982) note that based on the few studies on the growth of fish, the joint effects of toxicants has been consistently less than additive. The lack of additivity in growth effects on fish observed by Alabaster and Lloyd led EPA to recommend that one should not assume additivity when estimating the chronic toxicity of chemical mixtures for NPDES purposes, although no alternative model for handling the joint chronic toxicity of mixtures was offered (U.S. EPA, 1991; p. 24). Interestingly, in the absence of information to do otherwise, EPA recommends (in the same document) that one assume additivity of chronic toxicity when modeling the joint toxicity of effluents from multiple dischargers (U.S. EPA. 1991; p. 86). From another review of the topic, Suter (1993) states that several studies indicate the mode of joint toxicity of the mixture (additivity, antagonism) depends on the exposure duration and the response being measured.

In conclusion, the above evidence suggests that summing the individual ecological hazard indicators across chemicals (e.g., assuming additivity of hazard) is reasonable by several accounts. First, logic suggests greater hazard associated with numerous, equally toxic chemicals compared to one or few. Second, experimental evidences supports this assumption for acute toxicity of chemical mixtures and thus, directly supports the summation of ecological indicators that are based on acute toxicity data. Finally, while limited evidence suggest that the mode of joint chronic toxicity may be less than additive based on one endpoint for fish, and somewhat duration and response-specific in other cases, the assumption of additivity of ecological hazard (i.e., summing chronic ecological hazard indicators) is not unreasonable given the lack of consistent basis or model to do otherwise.

#### References:

Suter, G.W., II. 1994. *Ecological Risk Assessment*. Lewis Publishers, Chelsea, Michigan.

U.S. EPA. 1991. *Technical Support Document for Water Quality-Based Toxics Control*. EPA/505/2-90-001. Office of Water. Washington, D.C.

Alabaster, J., and R. Lloyd (eds). 1982. *Water Quality Criteria for Fish*. 2nd Ed. Butterworths, London.

## VI. OTHER ISSUES

### **Issue: Technical Questions and Corrections About Use of TRI Data**

#### **Agency (Regions) comments:**

The model does not take into account weekly fluctuations of the TRI database, due to open revision policy. The method should use data from the two week period when TRI data is frozen by Headquarters. (p)

*Response: The revised methodology documentation now states that data from this two-week period is used and that the data from all previous years is updated and recomputed at the same time.*

Before any methodology such as this can be applied to TRI data, the Agency needs to show that the changes being reported by covered facilities are in fact real changes in releases and offsite transfers of the chemicals. The statute only requires reasonable estimates, and EPA has interpreted that to be release amounts reported to two significant digits. Moreover, the most release estimates made by companies who report under TRI are based not upon monitoring data, but on emission factors, best engineering judgment and other estimating tools. There is no requirement that a facility use the same technique/process to estimate its releases from year to year. Thus, reductions in releases shown for a particular facility over time may simply be due to changes in its technique for developing "reasonable estimates," and not based upon actual reductions in releases of chemicals to the environment or transfers offsite. In order to demonstrate the majority of reductions in release and offsite transfer numbers are real reduction in emissions, the Agency would have to increase its efforts to review facilities' data through Regional data quality inspections, state grant data quality surveys and contractor site visits. (r)

*Response: As noted previously, the Agency is currently engaged in a data quality survey of individual facilities in several industry sectors.*

We urge you to have the Information Management Division rerun your lists of chemical in Tables 1-4. On a spot check of Table 1, we found two chemicals, incorrectly listed as having no reports. One of these is dicofol, for which an incorrect CAS number is given. If the searcher had looked under the correct CAS number 135-32-2, they would have found a number of reports for this chemical. The other we checked was Cupferron, CAS number 135-20-6. There were also reports for this chemical. On Table 4, picric acid is shown as a chemical with TRI releases of less than 1,000 pounds across all media. In 1990, the Dupont Victoria Site injected 34,930 pounds of picric acid underground at their facility. The Information Management Division has extensive experience with the TRIS database and would be able to rerun these tables for you with much improved accuracy. (r)

*Response: Additional effort has been devoted to performing QA on the data used.*

We would like to have more explanation in this document about how you developed Tables 1-3. It appears that you only considered direct environmental releases. We would prefer to have you combine both releases and offsite transfers when reviewing a chemical for inclusion in any of these tables. Figure 1 on page 4 supports this recommendation, since it clearly show that releases to air, groundwater and surface water can occur from chemicals transferred offsite to POTW's and other treatment and disposal facilities. Starting with the 1991 reporting year, the Agency will receive data on amounts of chemicals sent to recyclers and energy recovery units. These additional types of offsite transfers can also have direct environmental releases. We recommend that Criteria 2-4 be revised to include both direct releases and offsite transfers. (r)

*Response: The method no longer uses the criteria mentioned above for excluding chemicals. All chemicals are included unless there are no toxicological or physicochemical data available to run the model.*

#### **Industry comments:**

With regard to the baseline year selected for the program, the year chosen should be early, representative, and contain adequate data. We believe that 1988 might be a better choice of a benchmark year than 1989, since several other EPA programs have used 1988, including the 33/50 program. (v)

In the methodology report, a baseline year was discussed for comparison. We suggest 1988 SARA 313 reporting year to be the baseline year because it would be consistent with the EPA 33/50 program. (w)

*Response to above comments: The Indicators Work Group has adopted this suggestion of selecting 1988 as the base year for the Indicators. For trends analysis, the user can select a normalization of the data to this base year. However, the computer model is now structured so that all years of TRI data, or any subset of years, are available for analysis. Any year can be compared to any other year (with the exception of 1987, which is not included).*

#### **Issue: Environmental Modeling Assumptions/Approaches**

##### **Peer Reviewers comments:**

I think the assumption that "modeling equals bias" is less tenable for exposure modeling than for demographic modeling. Accordingly, I would urge you to only use a total of one order of magnitude (two "factors of 3") to distinguish site-specific from other types of exposure estimates. (a)

*Response: In response to this comment, only a single order of magnitude separates the highest from lowest exposure category.*

Your discussion starts out including environmental transport and fate on line 6, but by the last line you are dealing just with toxicity and exposure potential. Environmental fate should include persistence in the environment and transformation into other chemicals. (b)

*Response: Persistence is included as part of the aquatic toxicity score. Tracking the transformation of chemicals in the environment was deemed too complex for the purposes of the Indicator model; however, degradation is considered.*

Figure FLOW has too much detail on intermedia transfers (often minor) and not enough recognition on transformation and fate. (b)

*Response: Fate is included in the model: the figure simply shows which fate models are used to describe different kinds of releases, including intermedia transfers. Furthermore, "intermedia" transfers are quite significant, such as transfers to POTWs.*

"However, even generic modeling has significant advantages." Hooray for this insight, and please follow it where it leads. You need dose information in order to mesh the toxicity and exposure aspects of the methodology. (b)

*Response: The exposure potential component of the Indicator addresses the issue of meshing toxicity and exposure.*

Be very careful in modeling the release of chemicals into groundwater from RCRA nonhazardous landfill disposal units. In many cases you will need to model soil-chemical contaminant interactions. Get a geochemist to help you. (b)

*Response: The groundwater evaluation is done generically, relying on a national-level Monte Carlo analysis performed by OSW to estimate the transport of chemicals from a nonhazardous waste landfill to a nearby well. Given the effort and greater data requirements any site-specific modeling would require, the generic approach was deemed adequate for the Indicators effort.*

.. how are background exposures handled (i.e., shouldn't weight of threshold pollutant depend upon whether lots of other pollutants with same property are in the area? (c)

*Response: Background exposures are not considered because there is no practical way with the limited data to estimate these for **all** chemicals and all media at every TRI release site. The impacts of releases of threshold-acting toxicants should be considered as the ratio of exposure contributed by the source to the threshold level. The closer the exposure is to the threshold, the less margin of safety exists when accounting for background exposures. The Indicators investigate incremental risk in the absence of site-specific data. The methodology assumes that all background exposures (from sources other than TRI facilities and for non-TRI chemicals) at TRI release sites are equal in order to make such comparisons. Further analysis of higher risk-related situations should attempt to elucidate such site-specific information.*

**Agency (HQ) comments:**

It is logical to consider any well-known toxic products of degradation when evaluating the toxicity of a chemical. Perhaps an additional safety factor should be added when chemicals are known to break down into more toxic substances. It would probably not be worthwhile at this time to further complicate the modeling process by trying to model the rate and transport of degradation products as well as of the initial chemical. (h)

*Response: The method could incorporate such factors if data were readily available.*

(Calculating concentration of TRI chemicals in landfill waste). Using  $M_w$  as the divisor in landfill concentration may underestimate the concentration of the TRI chemical, since the landfill may include some chemicals from sources other than TRI facilities. (h)

*Response: Actually using  $M_c$  as the numerator may lead to an underestimate of the concentrations of the TRI chemical in the waste if there are other sources of the chemical in the landfill. This has been noted in the document.*

(Calculating concentration of TRI chemicals in landfill waste). By using constant 1986 values for annual waste volumes, this indicator could mask changes in waste concentration (and the associated risk) over time. (h)

How will the leachate concentration equation be adjusted for surface impoundments? (h)

*Response: To simplify the calculations, leachate concentrations will be calculated for surface impoundments in the same manner as leachate for landfills. While this introduces some inaccuracy, it is sufficient for the purposes of an indicator.*

When calculating human exposure to groundwater, this indicator estimates the exposed population as the number of well water drinkers within 4 km<sup>2</sup> of the landfill site. This seems overly conservative for three reasons. First, groundwater does not move in 360 degrees. The use of a 90 degree slice or other portion of the total area surrounding the site should be considered. Second, groundwater often moves very slowly and the masses of chemicals are limited. Exposures are not likely to occur 4 km from the landfill site. Finally, populations drinking from private wells are likely to be small. Since the indicator's scoring scheme ranks exposures from 0-999 as 1000, the well water drinker's population is likely to be overestimated. (h)

*Response: The method looks at populations within one kilometer from the facility. However, since the direction of groundwater flow is not known, the entire population in this radius is considered to be a potential exposed population. This will overestimate the population exposed, but choosing a 90 degree slice in a random direction may underestimate the population. For the purposes of the indicator, using the entire population was deemed adequate.*

The assumed size of the landfill unit is an issue, since it will affect fate and transport of the chemicals. (h)

*Response: The assumed size of the unit is an issue; the Indicators Work Group used the best available data to estimate this value.*

How will this indicator treat risks to populations served by municipal water systems, given that the Safe Drinking Water Act controls are in place? (h)

*Response: The effects of treatment are ignored, and this will tend to inflate the exposure estimates. For this reason, this exposure pathway is assigned to uncertainty category B; that is, the exposure potential estimates are decreased by a factor of five to account for this uncertainty.*

(Assignment of a chemical release to Subtitle D or Subtitle C land disposal unit). A facility with a RCRA I.D. number may have both Subtitle C and Subtitle D units. Thus there is the possibility that a TRI chemical at a Subtitle C facility could be going to a D unit. (h)

Although releases from Subtitle C unit could be negligible in comparison with those from a Subtitle D unit, the amount released from a Subtitle C unit is not always zero. (h)

*Response: It is true that the waste could be going to a D unit even if the facility has an ID number. However, the alternative is to assume that all waste is going to D units, and this was believed to be overly conservative. It is also true that risks from a C facility aren't always zero; however they are assumed to be low enough to exclude from the analysis.*

The paragraph that discusses dilution and attenuation of chemicals through ground water from a leaky landfill - this paragraph should be modified to include clarification that the EPA VHS and CML models used in the TCLP rulemaking neglect the effects of pumping wells (e.g., wells that would be pumped for drinking water). Note that the intent of VHS and CML models is to evaluate mobility of chemicals and not to determine dose. Because the TRI method does use the model results to calculate dose, the text should be clear that the dose calculation is oversimplistic because it neglects the effects of pumping wells. The effect of a pumping well could increase the groundwater velocities in the vicinity of the well and thus affect attenuation or dilution. (f)

*Response: The language in the methodology has been modified to reflect uncertainty.*

Why are we devising scores for underground injection disposal - it seems redundant. Doesn't OSW or Water regulate this practice? At any rate, UIC is regulated and is not supposed to result in exposure. (j)

*Response: Underground injection is not modeled. However, the Indicators do report the quantities of chemicals disposed of through this method. See Appendix E of the methodology for a proposal of an alternate method of addressing underground injection wells in the Indicators.*

The Methodology presents a reasonable overall approach for deriving a risk index for TRI releases. Our main concern is the generation of risk terms for scenarios where virtually the entire exposure calculation is based on default assumptions. Moreover, in many cases we may not even know whether the default assumptions are at all reasonable for the case at hand. This concern applies to most land disposal scenarios and to all generation of the waste at a TRI facility. For example, the received waste from a TRI facility would be highly speculative if the only hard data

available are from the TRI report. In cases where we can do a reasonable risk calculation, such as for direct air and water releases, we should by all means do so. In other cases, a risk calculation may be so highly speculative as to be meaningless. Furthermore, the generation of numbers that do not appear to be much better than wild guesses is bound to be controversial and could undermine the credibility of the entire system. (j)

*Response: It must be reiterated that this is an indicator, not a risk assessment. Conservative assumptions are adequate for the purposes of ranking chemicals and industries in a screening-level exercise.*

**Agency (Regions) comments:**

We strongly disagree with methodology's assumption to disregard degradation products since such products can have lasting health effects. (m)

*Response: The method might be able to incorporate data on degradation products if such data were readily available, and transformation rates and the toxicity associated with the degradation products were known.*

Modeling expertise does exist and should be available to EPA to include all factors in the calculation of this indicator. (m)

*Response: The reviewer did not cite any relevant sources for obtaining such information.*

It would be an excellent idea to embark upon a project to match publicly owned treatment works reported as offsite transfer locations in TRI with their National Pollutant Discharge Elimination System permit number (page 61). (r)

*Response: The project staff agrees that this would be a good idea, but such a project is outside our current scope.*

If underground injection releases were intentionally eliminated from consideration in these [release] tables, we disagree with that decision. It may be difficult to do an exposure analysis of underground injection releases. However, that doesn't mean we should drop them from consideration. (r)

*Response: Currently, underground injection is not included in the model because of the difficulty in hydrogeological modeling of potential exposure. The model has been modified to report pounds of chemicals sent to underground injection. The project staff are investigating alternatives (see Appendix E of the methodology).*

## **Environmental Groups comments:**

The exposure rankings are not likely to become a useful community evaluation tool in their current form. The current rankings embody many assumptions that few in the non-profit environmental community will agree with. In particular, it is assumed that there is no exposure from waste sent to underground injection wells, hazardous waste landfills, or hazardous waste incinerators. This ignores any exposure to workers at the facility that produces the waste. It is also ignores exposure due to accidental or fugitive escapes of the waste if it is transported outside of the facility. Lastly, there are definite exposures even from "properly designed" hazardous waste landfills and incinerators. For instance, incinerators emit both uncombusted waste and products of incomplete combustion (PICs) as well as residues, especially metals, in ash. Because of the assumptions of no exposure from these facilities, implicitly giving them a "clean bill of health," few in the environmental community will be able to support the Environmental Indicators as currently proposed even without any other objections to the draft report. (y)

*Response: These objections to the report still hold true. The Work Group made the explicit policy choice to leave out these exposure sources because (a) it is unclear how to model potential exposure from UI facilities (b) the risks from well-designed RCRA facilities are likely to be very low and (c) the relationship between incinerator feedstock concentrations and subsequent emissions of PICs is not well characterized. Of course, the Indicators do not track worker exposures. However, the model does still permit tracking of the pounds of individual chemicals that are disposed in these facilities. Furthermore, these pounds can be combined with the toxicity scores to yield a hazard-only indicator score - an alternative addressed in Appendix E of the methodology but not yet implemented in the model.*

Some national environmental groups might want to modify these exposure rankings and use them for their own purposes. After all, some of the rankings by environmental media embody useful data in themselves, such as which chemicals are "pass-through chemicals" and which are not in a POTW, or the data on relative rates at which chemicals settle out of the air. In order for the rankings to be useful for these purposes, they would have to be medium and chemical-specific without being facility-specific. Of course, these data are to some extent already available independently of the "TRI Environmental Indicators" project. Yet EPA would be providing a service to researchers with fewer resources by collating and checking these data. (y)

*Response: There have been numerous improvements in the computerized model of the Indicators. It is easy to aggregate data of various sorts beyond the facility level to look at specific media, chemicals, and types of releases.*

If EPA does intend to publish exposure rankings, generic models using facility-specific data should not be used. These models add an element of computational complexity that makes it impossible for environmentalists to re-calculate their own factors if they disagree with the underlying assumptions used (few environmental group have the computer resources to calculate an ISCLT generic air model for every TRI facility with air releases, for instance). It is difficult to believe that the models add any degree of accuracy that is not outweighed by the unknown factors underlying the analysis. For example, the uncertainty among different exposure pathways is taken into account by multiplying the results of different models by factors ranging over two orders of magnitude. When factors of 100 are being added due to uncertainty, it seems that facility-specific

differences in the results of the models (such as differences in local wind speed) would make little difference. Even some of the basic toxicity rankings are divided by a factor of ten due to uncertainty. In these conditions, it would seem better to assign exposure factors based only on the intrinsic properties of each chemical for each medium without modeling. This would remove the computationally difficult part of these calculations. If the modeling must be used, facility-dependent factors (local wind speed, etc.) should be eliminated and replaced by generic factors in the same way that true facility-specific data (such as stack height) have been. The models could be averaged and reduced to specific numeric factors for each chemical-medium combination that would at least not be facility-dependent. (y)

*Response: One benefit of modeling rather than using simple physicochemical properties for exposure scoring is that it allows comparison of exposure potential across media. Facility-specific modeling also allows for the interaction between environmental fate and geographic/demographic information.*

#### **Industry comments:**

EPA should use realistic assumptions to calculate the amounts actually released to the environment. On page viii, the document states that air releases from incinerators will be calculated using removal efficiencies for nonhazardous waste incinerators. In many cases, wastes are sent to hazardous waste incinerators. In order to calculate a meaningful measure of the impacts of chemical releases, it is important for EPA to use accurate assumptions about both the location and amount of releases to the environment. (s)

The statement is made that releases from incineration are to be modeled using "removal efficiencies for nonhazardous waste incineration." Obviously, this makes no sense if the offsite transfer is made to a licensed RCRA hazardous waste incinerator as is the case for the majority of TRI chemicals. The removal efficiency used should reflect the actual treatment method, in most cases RCRA hazardous waste incinerators. (t)

*Response to above comments: The method only assumes that wastes are transferred to a nonhazardous incinerator if there is no RCRA ID number given for the treatment facility ((i.e., it is assumed that facilities without RCRA ID codes are nonhazardous facilities). Wastes transferred to a hazardous waste disposal facility are assumed to pose minimal risk and are not modeled. However, the pounds of chemicals sent to such a facility can be tracked.*

Another concern of the draft report and proposed methodology is the use of measures for releases that are neither realistic nor based on accurate assumptions. In developing environmental indicators, the Agency should focus on actual releases to the environment. Any other measure is meaningless, yet in the document EPA has counted offsite transfers against the point of origin, instead of the point of release. The only useful measure of the environmental impact must be made at the site where a release occurs, which is at the treatment or disposal facility. (s)

It should be recognized that transfers to offsite locations, such as through a sewage system to a POTW, or tank truck transfers to treatment and disposal facilities, have no environmental impact relative to the site of origin. If the impact of the treatment process, for example wastewater treatment or incinerator is desired, a separate indicator should be developed for the

treatment facility. For example, Section V Mechanics of Estimating Exposure, (figure) shows arrows from offsite incineration and POTW transfers back to the reporting facility to be modeled as being disposed to the air or into surface waters. The concept of charging some portion of an offsite transfer back to the originating facility is totally unrealistic. In the case of the POTW, it may be located many miles distant from the discharging facility and may discharge to a different drainage basin. Many POTWs discharge to waters which are not drinking water sources. In the case of the offsite incinerator, it may be located hundreds of miles from the originating facility. If the EPA is concerned with human health and environmental effects associated with POTWs and waste disposal facilities, it should mandate these facilities to provide the required information and determine the impact on the area in which they are located. (t)

*Response: These two commentators are mistaken. The Indicators model the impact of offsite transfers at the treatment/disposal facility, not the TRI facility. The figure draws arrows not to the origin facility, but to the type of modeling that is done with the offsite release (such as groundwater or volatilization modeling). The text clarifies this relationship. However, the present model does add this off-site score to the reporting facility. The Environmental Justice Module of the Indicators will address such off-site transfers as cumulative risk-related impacts on the **exposed** population.*

Since these data may be interpreted as an exposure index, the assumptions made in the model are very important. For example, for an air emission source, a single exposure score will be assigned to all the population residing within a 10 km x 10 km area adjacent to the source. It is then very important that the median exposure in that population be calculated, rather than the maximum exposure in that population sector. Otherwise data will be misinterpreted as showing a much higher exposure than actually exists. (u)

*Response: The calculation uses the estimated concentration in each of the cells, combined with the population in the cell to calculate the exposure weight for that cell. The facility score is then the sum of these cell-level scores.*

The document states that the scores reflect impacts of releases/transfers, but do not provide quantitative estimates of risk. That is correct, however, it would not be difficult to back calculate order of magnitude risk estimates. Therefore, the assumptions and methodologies have to be carefully thought out, estimating average exposures rather than worst-case estimates. This is especially the case for situations where the data can be disaggregated, and exposures "back calculated" out of the indices. Exposure scenarios reflecting real-world scenarios, to the extent possible, will prevent calculating exaggerated exposure numbers. (u)

*Response: Where possible, average, realistic scenarios are used in the analysis. In cases where conservative assumptions are used (e.g., the groundwater model) the estimates are decreased by a factor of five or ten to account for some of this conservatism. It is important to note that the actual values derived by the indicator are unitless scores that only have meaning in comparison to each other; therefore, as long as the conservative assumptions affect the chemicals in the same direction and roughly the same magnitude, the effect on comparisons among chemicals should be minimized.*

The methodology does not explain how EPA will use the information in section 6 (transfers to off-site management, waste broker, or unknown for reporting years 1987 through 1990). If wastes are shipped to a waste broker or an other off-site location, the method of waste handling is not included in the SARA 313 report. Without a proper waste handling method, chemical exposure potential can not be determined. We suggest, if the waste handling method is conducted at a RCRA facility, the model should assume no further risk. This would be a similar assumption of handling RCRA waste that is outlined on page 66 of the EPA TRI Environmental Indicator. (w)

*Response: The method assigns offsite waste handling to a RCRA facility if a RCRA ID number is indicated for the offsite location.*

### **Issue: Aggregate Population Risk (Versus MEI Risk)**

#### **Peer Reviewers comments:**

Acknowledge that the entire methodology is very much oriented towards the "body count" axis of the "population risk versus individual risk" dichotomy EPA programs continually grapple with. For example, all of the early discussions of chemicals to consider excluding reflect an orientation implicit in the entire report-- that the size of a risk is determined by the size of the population consequences, not by any consideration of unacceptably high individual risks (excluding on the basis of "low" release values, for instance, ignores the possible importance of small quantities whose only impact is on small populations). The judgment to count all exposed populations as containing 1000 persons or more is a sensitive bow towards the urban-rural aspect of this dichotomy, but Abt might consider some more direct alternatives, such as: (a) use existing and forthcoming data (such as that which OAR is collecting under the Clean Air Act Amendments) to estimate the distance from each TRI facility to the nearest residence, well point, or surface water intake, and then compute a separate "MEI indicator," such as the number of facilities where one or more people are exposed to cancer risks above (say)  $10^{-3}$  or to concentrations of "threshold" toxins" above the RfD; (b) with additional data or estimates, EPA could compute an individual-risk-based indicator that incorporated population size, such as the number of persons facing cancer risks above  $10^{-3}$  and/or exposures above the RfD; or (c) a more arduous, but perhaps highly valuable supplementary indicator, would entail modifying the previous suggestion to yield an indicator of the number of persons facing either cancer or non-cancer risks above equivalent de minimis levels (perhaps  $10^{-6}$  for cancer and 1/100 the RfD for any [or 1/10 the "Hazard Index" for the sum total of all] exposure to threshold agents). The great advantage of this type of measure is that it would allow for the combination of cancer and non-cancer effect indicators without any implicit "severity weighting" of the actual health endpoint, by equating the risks somewhere at the "low end" of each scale only. Such an indicator would, of course, be highly inclusive and would not by itself identify progress in ameliorating serious individual or population exposure problems, but it would allow for tracking of progress in eliminating individuals or populations from any concern whatever, even if the baseline number which falls into this category is relatively small [see, if interested, the literature on risk perception and the "Russian Roulette" analogy, which identifies a perceptual premium on removing the hypothetical "last bullet" from the chamber when the baseline risk is already "small"].

Whatever [EPA] decides to recommend in this regard, I hope the spirit of my suggestions emerges: that a single chronic health indicator may be insufficiently rich to capture the various dichotomies (urban/rural, bodycount/MEI, cancer/noncancer) that pervade these issues. Just as "Consumer Reports" gives 4 or 5 colored indicators for the strengths and weaknesses of each product it evaluates (e.g., reliability, safety, fuel economy, handling, etc., for autos), so should EPA consider presenting multiple indicators and allowing each citizen to glean the information he or she deems most salient. (a)

*Response: The text of the methodology now specifically acknowledges a population-risk orientation. The other Indicators described in this suggestion have not been implemented; however, the computer algorithm could be modified to construct such indicators if these were desired. Also, the Indicators will no longer represent just one national number, since subsets of the TRI reporting universe may be investigated.*

Does the treatment of rural subpopulations effectively give greater per capita weight to sparse populations than for dense populations? Please defend this decision more extensively. (c)

*Response: Yes, the treatment of rural populations does give greater weight per capita to these populations. The Work Group decided to assign a minimum value of 1000 to rural populations so as to insure adequate representation for sparsely populated areas.*

Shouldn't the user be allowed to provide extra weight if the exposed population is more than, say, 10% nonwhite? Such data are probably readily available by geographic region. This is a good opportunity to make a statement about the need for consideration of environmental equity. (c)

*Response: An Environmental Justice Module of the Indicators is currently being developed.*

#### **Agency (HQ) comments:**

The aggregate population approach (proposed for the TRI Indicator) is consistent with the OGW/DW approach and appears sufficient since the TRI indicator is not attempting to reflect true risk assessment. (f)

While an MEI-based indicator would be useful for determining possible higher exposures of the chemicals to individuals in rare exposure circumstances, it would not be representative of exposure nation wide. MEI indicators may not capture real world actions to reduce worst-case sites and chemicals. Therefore the presentation of MEI indicator should be properly caveated. In evaluating the use of MEI analysis, it might be useful to consider the directive of the Hank Habicht memo on risk assessment uncertainty. Again, perhaps it would be best to run some sample data or simulations on the main indicator in order to see if a MEI sub-indicator is warranted. One alternative to the MEI indicator, which might be considered in the future, would be an MEP (maximum exposed sub-population) indicator for sensitive or highly-exposed populations of concern. This indicator could eventually tie into the environmental equity issue. (h)

*Response to above comments: No, MEI indicator is currently planned, but an Environmental Justice Module of the Indicators is currently being developed.*

(Population Size Weights) Rounding up to the nearest 100 will probably have negligible effect on the result. However, the stated rationale for it is questionable. Value judgments related to social policy concerns should not influence the indicator. This compromises its status as an objective indicator of risk. While it is entirely legitimate and appropriate to consider individual risk as well as population risk in the indicator, this should be done directly rather than by introducing some kind of modifying factor into population risk estimates. OPPT's standard methods for evaluating ambient air exposures to TRI releases allow for estimating of individual risks as well as population risks. (j)

**Environmental Groups comments:**

When all of these factors are determined and multiplied together, the result according to the draft is a “relative risk ranking” specific to each chemical-medium-facility combination. Even if we discount problems with this ranking caused by uncertainties in the data or incorrect assumptions, the result lends itself to misuse. Since all facilities are by this definition “lower risk” if fewer people live near them, this scheme encourages buyouts of nearby home or siting of facilities in rural areas rather than true risk reduction actions. It de-emphasizes toxic use reduction and source reduction and encourages actions such as shifting waste from one release medium to another. (y)

*Response: As a population risk-based Indicator, by definition risk will be lower if fewer people are exposed. This is inherent to the nature of the Indicators. However, the Indicators do track trends and permit the user to select the quantities of emissions as an analysis parameter. This can be used to identify whether chemical emissions are being shifted between environmental media.*

**Issue: TRI Environmental Indicators and GIS**

**Agency (HQ) comments:**

Funds have not been made available to connect TRI Indicator to any GIS. There is strong concern that continuing to move forward (with the indicator) without doing so would be an important mistake. Connecting the indicator to ARC/INFO would provide the opportunity to relate indicator results geographically to locations of sensitive human populations, ecological features, observed health and ecological impacts and locations of other impacts. [The Environmental Results Branch] is interested in raising funds from a variety of sources to support the programming needed to connect the indicator to ARC/INFO. (e)

*Response: The analyses available using the Environmental Justice Module of the Indicators can be examined directly by the user at limited scale or the results can be linked with a GIS model.*

## **Issue: Calculation of the TRI Environmental Indicators**

### **Industry comments:**

If the Agency desires a strict indication of annual change, the proposed method of comparing each year's emissions with the baseline year will provide those results. If, however, the objective is to more accurately identify trends, we suggest that the data would be significantly improved by using a rolling average of three years. For example, if the baseline year is 1988, the first-time comparison might average emissions data from 1989, 1990, and 1991; the second-time comparison from 1990, 1991, and 1992; and so on. This method has the advantage of smoothing out fluctuations which inevitably occur from year to year, and which result from regulatory changes to the data reportable on the TRI, or various unusual occurrences at a facility. For example, a refinery's emissions might increase during any given year as a result of one-time remediation activities. Although remediation provides long-term environmental benefits, the short-term effect is to increase emissions, which appears as a negative event with respect to the TRI and the EI. Both approaches are valid ones; we are simply suggesting that EPA carefully consider which will best serve their purpose and communicate the implications of their choice to those who will use the EI data. (v)

*Response: The current method still uses the single year approach as the most direct way of reflecting changes in relative risk-based impacts.*

## **Issue: Normalizing the TRI Environmental Indicators**

### **Peer Reviewers comments:**

Consider "normalizing" the indicator(s) over time by recomputing back into the past rather than by "starting over" or adding new indicators. As chemicals are added to the TRI list, it should be possible to make a rough estimate of what their release characteristics were in prior years. In this way, continuity could be restored without a proliferation of indicators. (a)

*Response: OPPT disagrees that it would be possible to make sufficiently accurate estimates of past years' emissions in order to recalculate the Indicator values, especially on a facility-specific basis. Even if such estimates could be made, the associated uncertainties would put into question the validity of any conclusions made from analyses of such estimated data.*

I think you do have to have an algorithm for renormalizing the indicators as the list of chemicals and the information base for the chemicals changes from year to year. Explain what you do to renormalize and give the results with and without renormalization. See my illustrative example in the General Comments section. (b)

*Response: The method for normalizing is explained in the revised methodology documentation. Examples with and without normalization could be part of any reports describing Indicators results.*

## **Industry comments:**

How will EPA address additions/deletion to the chemical list and additional new reporting facilities? EPA is proposing to add 68 new chemicals and 2 chemical categories to the SARA 313 list. The baseline year for comparison will not include the new chemicals. EPA would not have a valid baseline year for comparison or establish trends because of the additional chemicals. We suggest EPA should not include the new chemicals indicators for three years. After three years, a new baseline could be established with sufficient data for comparisons. A similar result would occur when new facilities are added to the SARA 313 reports. (w)

*Response: The method discusses an approach where a new Indicator would be created when significant additions are made to the TRI roster. As part of the new Indicator, a sub-Indicator consisting of only the original set of chemicals would be separately tracked so that earlier years of the Indicator values could still be compared against new data.*

We recommend that the Agency consider in advance how corrections made to the TRI database will be accommodated in the EI methodology. As a result of these revisions, the TRI has become a dynamic system, quite different from the static database which was originally envisioned. As better data and more accurate assumptions lead to more precise measurements and estimates, facilities refine their emissions reporting for both current and past years. Correctness dictates that such changes should be reflected by revised EI values, and we suggest that the methodology should include a discussion of whether and how this will be accomplished. (v)

*Response: The Indicators will be rerun each year for all available years, during the annual TRI data freeze, so that changes in the reporting database can be incorporated.*

## **Issue: Expansion of the TRI Environmental Indicators**

### **Peer Reviewers comments:**

p.41 Having been involved in risk assessment activities on global warming, photochemical smog, stratospheric ozone, and acid rain, I found your willingness to try to expand TRI indicators into these realms a welcome bit of levity at the end of a long, complex report. As your discussion points out, the difficulties are enormous, because of the details of the processes that need to be modeled -- and massive efforts have been undertaken by the Agency in modeling risks in these four problem areas. You may wish to coordinate with the people in EPA that are working in these four areas, and from their knowledge, developing suitable indicators (or using the ones they already have developed) as part of a reporting system for strategic environmental planning. (b)

*Response: Appendix I of the methodology discusses these issues at length.*

# **Appendix**

## **Peer Reviewer's Comments**

March 29, 1992

Ms. Susan Egan Keane  
Abt Associates, Inc.  
4800 Montgomery Lane  
Suite 500  
Bethesda, MD 20814-5341

Dear Susan:

In addition to the comments I gave you verbally on February 28, I wanted to transmit the following comments on the "TRI Indicators Methodology" report (9 January 1992 draft). First and foremost, you and your colleagues are to be commended for preparing a particularly thorough, well-written, and thoughtful report. I don't intend to damn with faint praise by offering what compared to other consultants' reports and documents "written by EPA" I've read lately, yours reads like Shakespeare (and is generally well-reasoned to boot)! On the substantive front, I particularly commend Abt for making several important "judgment calls" in what I consider to be appropriate and sometimes novel ways: (1) striking a good balance between "full-blown" risk assessment orientation (which might arguably be too data-intensive) and a wholly ordinal scheme that fails to make use of information that does exist and that is too easily over-interpreted (I would argue, by the way, that the HRS used in the Superfund program is too much like the latter than the former); (2) eschewing the temptation to combine disparate health endpoints by the highly subjective and overly simplistic process of "severity weighting"; and (3) creatively accounting for uncertainties about which only the sign is presumed known (e.g., weighting factors on pp. 22-23 reflecting the quality of population data). Moreover, the tone of the document is admirably balanced, with appropriate warnings against over-interpreting the results given prominent status and yet not belabored to such an extent that the methodology comes across as arbitrary.

There are important aspects of the methodology that troubles me, however. Some of the more fundamental concerns I have relate to the following aspects:

- (1) The decision is made to equate exposure to carcinogens and noncarcinogens via a weighting scheme that equates, for example, a  $q_1^*$  value of 0.1 (the midpoint of the range 0.05-0.5) with an RfD of 0.001 (and  $q_1^* = 0.01$  equates to RfD = 0.01). The net effect of assuming that the two scales can be made proportional *and* of choosing the proportionality constant such that these equalities hold, is to equate the following *risks*: exposure at the RfD equals a cancer risk of  $10^{-4}$ ; exposure at 1/10 the RfD equals a cancer risk of  $10^{-5}$ ; exposure at 10 times the RfD equals a cancer risk of  $10^{-3}$ ; and so on. These are highly value-laden assumptions. The first of these three points of equality may make sense to the authors, but the 10,000:1 implicit weighting of the cancer:noncancer endpoints must at least be highlighted. And yet if this equality makes sense, it seems hard to justify some of the other equalities, either at the "high end" (exposure to 1000 times the RfD seems even worse than a cancer risk of  $10^{-1}$ ) or at the "low end" ( $10^{-6}$  risk may be *de minimus*, but according to some, exposure to 1/100 the RfD is trivial by definition since it is far below a threshold of effect). In addition to making much more prominent the

implicit judgments buried in this system, you might consider maintaining two separate health indicators (one for cancer, one for toxic effects) so that there would be no need to linearize a probable threshold phenomenon, or use an alternative weighting scheme that equalizes risks at the “low end” only (see below).

- (2) the enclosed articles from the *Journal of Policy Analysis and Management* express my concern that the EPA’s alphabetical system of carcinogen classification may be misleading, and is probably inappropriate for use in quantitative potency adjustment. If Abt does not want to abandon such adjustments entirely, I recommend it at least compress its three categories into two (one for Groups A, B1, and B2, and one for Group C), with only a single factor of 10 separating A and C, rather than the current proposal for a 100-fold factor. This would not only reduce what I think may be a false dichotomy, but would make the indicator more stable over time, as chemicals shuttling between B1 and B2 status would not affect this alternative indicator.
- (3) I think the assumption that “modeling equals bias” is less tenable for exposure modeling (pp. 20-21) than for demographic modeling (pp. 22-23). Accordingly, I would urge you to only use a total of one order of magnitude (two “factors of 3”) to distinguish site-specific from other types of exposure estimates.
- (4) As with risk estimation in general, I firmly believe that estimates such as “indicators” can be reduced to single point values, but only if these are conscious and consistent choices emerging for a quantitative uncertainty analysis (QUA) rather than from a black box containing a mixture of “average,” “conservative,” and “anti-conservative” procedures. I realize that a quantitative calculation of uncertainty may be daunting if done *routinely*, but I urge you to urge EPA to at least undertake state-of-the-art QUAs for a few illustrative cases within the TRI project. That way, an appreciation for the “noise” in the indicator can be communicated. More importantly, regulators and the public could be educated regarding which comparisons (e.g., longitudinal comparison of indicators over time, snapshots across regions, industrial sectors, etc.) involve the *compounding* of uncertainties and which probably involve the *canceling* of uncertainties. For example, suppose the indicator moves from “ $100 \pm 20$ ” to “ $80 \pm 15$ ” between 1995 and 1997. If the uncertainties are not strictly parallel, then I would argue the “noise” outweighs the “signal” of progress (the 1995 value might well have been 85, and the 1997 value might well have *risen* to 90). However, if a model QUA revealed that the sources of error were common to both measures in this comparison, then the apparent rank order might also be robust.
- (5) It may be outside of your assignment from EPA, but it should be obvious that even the most rudimentary uncertainty or sensitivity analysis of the kind referred to above will have to include attention to validating the TRI emissions data. This would include attention to the accuracy of those emissions that are reported and to

the compliance levels of the affected firms (in case under-reporting is a problem). Are you confident that EPA has these problems well in hand?

- (6) One “risk communication” issue — I think EPA needs some guidance in how to investigate and explain changes in TRI indicators over time. Depending on the large national trends that will not be tracked as part of EPA’s work on this particular project, changes in the TRI indicators could reflect real progress (or lack thereof) in reducing the volume of emissions (or in switching to less toxic pollutants), or they could reflect changes in national demography (people migrating toward or away from industrial facilities), changes in intranational or international competition (polluting firms relocating to other regions of the U.S. or to other countries), or the definitional changes in TRI referred to in Section VI of your report. In my opinion, EPA needs some guidance on how to investigate the causes of any “signal” it detects in the indicators, and in how to communicate to the public any mitigating factors that might explain changes in the indicators (or at least the qualitative uncertainties introduced by such factors).

To address some of these rather general concerns, I would offer a few other specific (and, I hope, constructive) suggestions for changing the methodology or supplementing it with other approaches:

- acknowledge that the entire methodology is very much oriented toward the “body count” axis of the “population risk *versus* individual risk” dichotomy EPA programs continually grapple with. For example, all of the early discussions of chemicals to consider excluding reflect an orientation implicit in the entire report -- that the size of a risk is determined by the size of the population consequences, not by any consideration of unacceptably high individual risks (excluding on the basis of “low” release values, for instance, ignores the possible importance of small quantities whose only target is on small populations). The judgment (on page 23) to count all exposed populations as containing 1000 persons or more is a sensitive bow toward the urban-rural aspect of this dichotomy, but Abt might consider some more direct alternatives, such as: (1) use existing and forthcoming data (such as that which OAR is collecting under the Clean Air Act Amendments) to estimate the distance from each TRI facility to the nearest residence, wellpoint, or surface water intake, and then compute a separate “MEI indicator,” such as the number of *facilities* where one or more people are exposed to cancer risks above (say)  $10^{-3}$  or to concentrations of “threshold toxins” above the RfD; (2) with additional data or estimates, EPA could compute an individual-risk-based indicator that incorporated population size, such as the number of *persons* facing cancer risks above  $10^{-3}$  and/or exposures above the RfD; or (3) a more arduous, but perhaps highly valuable supplementary indicator, would entail modifying the previous suggestion to yield an indicator of the number of persons facing *either* cancer or non-cancer risks above equivalent *de minimus* levels (perhaps  $10^{-6}$  for cancer and 1/100 the RfD for any [or 1/10 the “Hazard Index” for the sum total of all] exposure to threshold agents). The great advantage of this type of measure is that it would

allow for the combination of cancer and non-cancer effect indicators without any implicit “severity weighting” of the actual health endpoint, by equating the risks somewhere at the “low end” of each scale only. Such an indicator would, of course, be highly inclusive and would not by itself identify progress in ameliorating serious individual or population exposure problems, but it would allow for tracking of progress in eliminating individuals or populations from any concern whatever, even if the baseline number who fall into this category is relatively small [see, if interested, the literature on risk perception and the “Russian Roulette” analogy, which identifies a perceptual premium on removing the hypothetical “last bullet” from the chamber when the baseline risk is already “small”].

Whatever Abt decides to recommend in this regard, I hope the spirit of my suggestions emerges: that a single chronic health indicator may be insufficiently rich to capture the various dichotomies (urban/rural, bodycount/MEI, cancer/noncancer) that pervade these issues. Just as “Consumer Reports” gives 4 or 5 colored indicators for the strengths and weaknesses of each product it evaluates (e.g., reliability, safety, fuel economy, handling, etc., for autos), so should EPA consider presenting multiple indicators and allowing each citizen to glean the information he or she deems most salient.

- consider exploiting the statistical relationship between acute toxicity ( $LD_{50}$ ) and carcinogenic potency ( $TD_{50}$ ) to sidestep the problem of different endpoints and of potential carcinogens for which long-term bioassay data are not available. Although it is still controversial whether the observed correlation is artifactual or reflective of an underlying biologic mechanism (I can provide articles by Zeise, Crouch, Starr, Gold, etc., if you are interested), it may still be useful for decision-making and for providing incentives for data generation by the private sector.
- consider “normalizing” the indicator(s) over time by *recomputing* back into the past rather than by “starting over” or adding new indicators. As chemicals are added to the TRI list, it should be possible to make a rough estimate of what their release characteristics were in prior years. In this way, continuity could be restored without a proliferation of indicators.

Finally, I have a few minor technical comments:

- (1) on page 4, I would urge OTS to involve other agencies (especially NTP, NCTR, ATSDR) to ensure that all available toxicity data are assembled, not just that contained within EPA offices.
- (2) on page 5, I would simply mention that a cancer potency factor of 0.01 per mg/kg-day dose does not seem particularly “low” to me (exposures to such compounds would have to be less than 1  $\mu\text{g}/\text{kg}\text{-day}$  to yield excess risks less than  $10^{-5}$ ).

- (3) on “Table SELECT” (ff. page 6), I note that since all but 1 of the 317 chemicals are *either* released in non-zero quantities (and may thus present individual risks that are worthy of consideration) or are not classified with low toxicity, that this argues for no exclusions at all at the outset.
- (4) on page 13, I understand that Abt must recommend the use of available data and models, but in the case of metals speciation (as well as other “default” assumptions mentioned elsewhere in the report), language encouraging EPA to *improve* these defaults over time would be helpful.
- (5) in “Figure UpUp” (ff. page 15), the number “ $1 \times 10^{10}$ ” in the middle of the page is a typo — it should read “ $1 \times 10^9$ .”

Thank you for the opportunity to comment on this excellent draft report. I hope these observations are helpful.

Best regards,

Adam M. Finkel, Sc.D.

**Mtg with Adam on TRI, 2/29/92**

- stay away from too simple
- discard some of the convoluted for final report
- more readable
- does a disservice to say scheme
  - looks cheesier than it really is
  - not that qualitative
- numerical system → combine cancer risk scale + noncancer risk scale
- more than indicator
  - # people above level of concern
  - MEI orientation
  - exposure profile
  - census → nearest grid
  - \* ask Brad about finer divisions
- what to do about factors of 10
  - + letters for WOE
  - could be in group A because causes common cancer
  - As + Bs → ambiguity of animal
  - Cs → weird mech.

Problem with RfD vs data and uncertainty factor

A+B+C differ by probability if it is.

Uncertainty cuts in one direction

- is the exposure always lower
  - no incentive to gather more data
  - lower the diff between A, B, & C — factor diff
  - What would factor be? Compare to site-specific estimates for magnitude
  - argues for consistency in using uncertainty
- guidelines for when comparisons are robust and when they are not
- thought given to what are params can be grouped, which are different
- uncertainty
- signal/noise
- which trends are significant, which comparisons are legit
- comparisons of distributions → as many as for diff comparisons we need to make
- what kinds of comparisons will be done — how to interpret changes in scores
- # could go down because fewer people
- as new chemicals are added → how to normalize over time
- MEI
- # people above threshold

Page 2

→ put everything →

All chemicals on LD<sub>50</sub> scale

acute → linear correlation with potency or effect  
or 1/MTD

→ nothing should be excluded

February 10, 1992

Ms. Susan Egan Keane  
Abt Associates Inc.  
4800 Montgomery Lane, Suite 500  
Bethesda, MD 20814-5341

Dear Ms. Keane:

Enclosed is my review of the draft report, "TRI Indicators Methodology." Please contact me if you have questions or need further assistance or information.

Sincerely,

D. Warner North  
Principal and Senior Vice President

DWN:scm

Enclosure

**Review of TRI Indicators Methodology  
Draft of January 9, 1992**

**Contract 68-D9-0169, Work Assignment 3-2**

**Review by Warner North  
February 10, 1992**

**General Comments**

Potential Misuse of Quantitative Indices

The initial reaction of this reviewer was one of horror. The proposed TRI Indicators Methodology is an extremely complex and ambitious effort to summarize information about toxics in the environment into a single quantitative index, an equivalent to the Dow Jones Industrial Average (DJIA) as an indicator for the stock market. The reason for the horror is that this reviewer, by nature and training highly disposed toward quantitative methods, has had many negative experiences with the efforts by EPA and similar agencies to develop and use quantitative indices for problems that are inherently complex and uncertain.

All too often these indices are misused, because managers ascribe a precision to them that is inappropriate (and frequently never intended by the developers). Rather than promoting understanding of the important issues of science that should be central to environmental risk management decisions, the quantitative indices often suppress these issues. Managers just want to look at the numbers.

Much criticism has been written about EPA's use of cancer potency "unit risk," or  $q_1^*$  numbers. While the virtue of this practice is simplicity and consistency across a large number of carcinogenic chemicals, the vice is poor communication. While these single-number plausible upper bound numbers may be useful for crude screening, they do not indicate the uncertainties and scientific subtleties that ought to be considered in major decisions on regulation of important carcinogens. Moreover, even the important qualification that these  $q_1^*$  numbers are plausible upper bounds is often lost when cancer risk numbers are communicated to decision makers and to the public.

Cancer risk assessment is looked on by many within the Agency as a big success. I do not think this view is shared by Congress, which set up a special NAS/NRC Committee (which two of your three reviewers are on) to review EPA's risk assessment practices when they passed the Clean Air Amendments Act (CAAA). The scientific community's discomfort with quantitative risk assessment is also considerable. Former Administrator Lee Thomas feels that Congress had to go into MACT in the CAAA because quantitative risk assessment was not providing the needed basis for decisions on control of airborne carcinogens. The recent report, *Reducing Risk*, by the EPA Science Advisory Board, used qualitative means to rank risks and emphasized the need for expert scientific judgment as the basis for the ranking. The Health Subcommittee for this

report did not attempt any quantitative ranking among carcinogens, substances with other chronic health impacts, or acute toxicants.

The Superfund Program has tried a number of scoring procedures for ranking their sites, and then compared the results to the judgments of an expert panel of senior EPA personnel. The lack of agreement in the rankings was striking. I've been involved in several rounds of review of the Hazard Ranking System, and currently, in litigation challenging the HRS as non-responsive to the statutory mandate with respect to regulation of fossil fuel combustion wastes. These ash materials are disposed of in very large volume (millions of tons), so the low concentrations of toxic metals they contain (arsenic, cadmium, nickel, etc.) yield large quantities (tons). The short version of this story is that the HRS will list sites that EPA has spent millions of dollars studying, after which EPA then reported to Congress that these ash sites did not warrant regulation as hazardous under Subtitle C of RCRA. So why should they be listed as Superfund Sites? Because EPA has tried to use a single formula for site evaluation instead of more in-depth scientific evaluation using the available information.

So, I hope I have made my point, and I will end this diatribe against overreliance on quantitative methods. I am very concerned about potential misuse of the TRI Indicators. Your caveats and your description of the communication aspects of TRI need a great deal of expansion. I will return to this point later on.

#### Impressions of the TRI Indicators Methodology Draft

Having stated the concerns from my understanding of your assignment, I then went on to read the document and try to understand what you did. While on the budgeted effort I cannot hope to check your data or the details of your methods, I am impressed that you have worked hard to assemble the data sets, develop reasonable aggregation procedures and decision rules for what data to include, and that you have done a thorough search of other people's scoring systems.

My second impression, therefore, (following the first one of horror), is that your system could be valuable in support of strategic planning and the Agency's communication to the public, in support of "the Agency's desire to set priorities and shift resources to areas with the greatest opportunity to achieve health risk and environmental risk reductions." (p.1 of your draft).

You recognize the limitations of the TRI and other data sources that reflect multimedia trends in environmental contaminant releases. The methodology must build upon the data available. Unfortunately, this is a very large limitation because of the lack of data and unevenness in data quality. In particular, you face enormous difficulties in trying to assess what emissions imply for exposure to humans and environmental receptors. Without models and the data to drive them, source-to-receptor relationships can be described only in the most generic terms. Further, chemical substances transform into other substances that may be more toxic, less toxic, or non-toxic. How quickly such transformations occur may depend on specific characteristics of the medium — air, water, soil — that can vary in time and space and with the concentrations of other chemicals present. So even with very good data on emissions, it can be very difficult to assess the consequences of the emissions in terms of risk to health and the environment.

Given these difficulties, what does the reviewer suggest you do differently? It's a matter of focus.

1. Do not try to maximize the comprehensiveness — adding as many chemicals and industries as possible. Rather, **concentrate on doing a good job for a selection of the most important toxic substances of each risk class.** The DJIA does not include all the industrial stocks on the NYCE, but rather a selected set of important ones. The Background Section in the Statement of Work talks about “capturing a representative cross section of toxic releases.” That sounds like a good way to conceptualize what you are trying to do.

My impression from the “Six Months Study” and subsequent attempts to assess (upper bound) national cancer risk from hazardous air pollutants is that most of the risk comes from a small subset of the chemicals — 20 to 25 of the 189 on the CAAA list.

2. Don't just use the data in the data bases. **Utilize expert judgment** from the program offices inside the Agency and the scientific community outside the Agency to assure that what you are doing is sensible. In many cases you might pick up and use existing indicators and quantitative summaries — for example, the cancer risk estimates for air toxics done by the Air Office. If your numbers do not match their numbers and their judgments, you should find out why. Your computer implementation of the TRI indicators methodology should be sufficiently flexible so that methodology and data refinements can be readily accomplished. Trial use of your system and extensive interaction with various parts of the Agency (and the outside scientific community) should give you opportunities to make important improvements over the methodology and data you start with.

I personally feel you should focus more effort on transformation, fate, and persistence in your methodology. Intermedia transfers need to be included, but I think you could waste a lot of effort trying to be comprehensive. The biggest concerns for health and environmental risk are usually the chemicals that are persistent (PCB's; mercury) or that transform into more toxic chemicals (chlorinated solvents into vinyl chloride through microbial action in aquifers; methylation of mercury). Very toxic chemicals that quickly degrade into harmless ones will be of concern only near the sources. Your TRI system ought to tell you that one of the biggest sources of human health risk from drinking surface water is the chlorination process, which produces trihalomethanes and lots of other chlorinated organic compounds, many of which are poorly understood and some of which are probably quite dangerous. Another important area with poor emissions data is the runoff of agricultural chemicals into surface waters and percolation into ground water. Indoor air is an underrated area for risks, because of the extensive use of toxic chemicals in home and office products/equipment and in building materials. The Agency focuses on radon and asbestos and largely ignores many other chemicals emitted in the indoor environment. Your system should indicate the need to shift priorities and resources in order to reduce human risk, even if some of the issues are outside EPA's statutory jurisdiction.

You should find lots of similar insights from a careful reading of the main report and the supporting subcommittee reports for *Reducing Risk*.

### Communication in Support of Strategic Planning

What I urge that you focus on is the use of the indicators for communication. I believe that was what Administrator Reilly intended when he called for the development of such indicators. One of Mr. Reilly's first actions as Administrator was to commission *Reducing Risk*, and in a recent issue of the *EPA Journal* (March/April 1991; Vol. 17, No. 2) you will find him calling for national debate and discussion on using risk as a basis for setting environmental protection priorities. You should see your efforts in this context.

I think it is a mistake to focus on one overall indicator for environmental risk. You mix apples, oranges, and all sorts of other fruits into a compote no one will find very satisfactory. It will be much better for you to focus on a "set of indicators," which is the term in the Statement of Work. Then for each one, you can discuss changes in a way that will be readily meaningful for the public.

Illustrative example. The index of toxic air contaminants is down 25% for 1993 compared to the 1992 level, assuming the same evaluation of toxic substance emissions for the two years. However, of the 189 chemicals on the CAAA list, three more were found to be carcinogens as a result of recent animal testing by the National Toxicology Program, and two carcinogens had potencies reduced as the result of research on the biological mechanisms by which chemicals cause cancer in animals and humans. Changes in the evaluation methodology corresponding to these changes in the knowledge base for toxic air contaminants result in an even larger decrease in the indicator, from 25% to 28%. Using EPA's standard estimates for assessing cancer risks, the revised 1993 index indicates that up to 1440 animal cancer cases might be attributed to toxic air contaminants, compared to the 1992 index estimate of up to 2000 animal cancer cases.

Similar paragraph descriptions might be written about each of the other indicators in the set. Where an indicator corresponds to a program office in the Agency, the paragraph could provide a useful summary of how this program office is doing, by relating changes in the indicator to program activities affecting various source categories. The program office will clearly have strong interest in what this summary says. Where there is no program office, as for example, with indoor air, the paragraph description would be a good indication to Congress — and the public — of the importance (or lack thereof) of shifting more resources into this area.

A summary of which indicators moved up and which moved down would be instructive. Also, looking at the pattern of the changes could be useful, especially in comparison to where the Agency is placing its resources. The change in a weighted sum of all the indices might be least meaningful for strategic planning.

## Specific Comments on the Text

- p.1 You go from “indicators,” plural, as what the Administrator wanted in paragraph 1 to the singular, “indicator,” in paragraph 2. As discussed above, I think this change in emphasis to a single indicator is a poor idea. The discussion seems to stay with the singular from then on: “the methodology and data sources used to develop the TRI indicator” (page 1, last two lines). In practice, you should be using at least two indicators, for health and for environmental effects.
- p.2 para 1. Your discussion starts out including environmental transport and fate on line 6, but by the last line you are dealing just with toxicity and exposure potential. Environmental fate should include persistence in the environment and transformation into other chemicals.
- p.2 para 2 and 3. Yes, site-specific risk assessment is hard and data-intensive. But by neglecting important aspects you can reach inappropriate conclusions. Do the risk assessment for “a representative cross section of toxic releases” (Statement of Work) at representative sites for releases into the appropriate media. Work with experts to find out how to do this properly. Then try to make the calculation simple and the results easy to communicate to the public.
- p.3 1st new para. Good to focus on four indicators, not one. Also, you are right to bound out spills and episodic releases. Make it clear that you are leaving out acute risks to health and the environment. To do this would require an enormous effort, given the nature of the time series data needed to do in order to do the job well. Don’t expect this time series data from future reporting under the TRI system.
- p.3 Section III, A. Go for chemicals representative of those posing the highest risks. Don’t limit yourselves only to the chemicals for which there is good release data. A valuable aspect of your system could be to go from observed concentrations posing high risk (in air, surface water, ground water, and soil), ask (the experts) where these concentrations might have come from, and then determine what additional data you need to get on emissions/releases in order to determine the source contributions more precisely.

Figure FLOW has too much detail on intermedia transfers (often minor) and not enough recognition and transformation and fate.

Table No-Rep. What are these chemicals used for? Is there a potential for human exposure except through direct ingestion of a food or drug? (Not likely for the food coloring agents.) Ask the experts or look up the chemical in a handbook. DBCP is an example of a pesticide no longer in use that caused very serious groundwater contamination, which still persists.

Criteria 2, 3, 4 and 5 are simple screening rules, but for some chemicals they could lead to important omissions. Follow your idea under criterion 4 of circulating the list to get more

data. Even small releases can lead to danger where there is high human exposure. Many proprietary chemicals will not be in IRIS, but OTS may have toxicological data on file under confidentiality restrictions. Low toxicity compounds can give high risks with high doses. Do you know how many people are exposed to high doses of acetone? toluene? ethylene glycol? If you do not, find a chemist in Abt Associates and ask him/her to explain what some of these chemicals in the Low-Tox list are used for.

Table Rep-Tox. You might inquire why dichlorobromethane and bromoform should definitely be in your system.

- p.6 Inclusion of facilities. I wonder if your decision to include all facilities makes sense. You will count lots of small ones, then ignore all those under the cutoff for reporting. Consider dry cleaners who emit perchloroethylene, or furniture makers who use paints and solvents. Will you get accurate estimates of total releases?
- p.7 Why do you use the acute indicator? I thought you decided to drop or defer it since the needed data on exposure over time are unavailable. See page 3 and comment on it above.
- 8-12 Weighting schemes should be approached with humility and great caution. Use of weight of evidence tables such as on p.8, 9, 10, 11 should be for crude ordinal ranking, not cardinal ranking. Simple severity times dose as at the bottom of p.12 may be a good, starting point for non-carcinogens (subchronic and chronic, not acute).
- p.13 Assuming that noncancer chronic health risk varies with the ratio of the dose to the Reference Dose will be highly misleading when this ratio is less than unity. You should be finding out how many people may be exposed to levels near or above the RfD.
- p.14  $ED_{10}$  and  $q_1^*$  are highly correlated, and one can be essentially inferred from the other.

The comparisons of vinyl chloride and benzene in Figures UPDOWN and UPUP are interesting and useful. You should have many other comparisons like these to illustrate how your schemes give results that are consistent with, or at variance with, common sense/expert judgment.

I think there will be many problems with your scheme for both carcinogens and non-carcinogens. Many of these you may be inheriting from previous EPA systems like HRS. What you get is an extremely crude score that may be useful some of the time. In other cases, look out. Try arsenic, a category A carcinogen that is one of the more common elements in the earth's crust. You will calculate very high potency scores for any water or soil in which there is detectable arsenic.

- p.17 "Applying such weights across categories of toxic endpoints would require a subjective evaluation of the relative severity of the health effects." You reject the difficult subjective evaluation in favor of using the most severe endpoint — regardless of the relation of RfD

to exposure. I predict a lot of your results will not make sense. Do some examples. How would you evaluate lead?

From this point I'm going to give you far less comments, to avoid being repetitive. You have generally followed what others at EPA have done in constructing scoring systems. The results of some merit for crude ordinal ranking, but for some important chemicals the systems may give misleading results.

- p.25 "However, even generic modeling has significant advantages." Hooray for this insight, and please follow it where it leads. You need dose information in order to mesh the toxicity and exposure aspects of the methodology.
- p.26 Be very careful in modeling the release of chemicals into groundwater from RCRA nonhazardous landfill disposal units. In many cases you will need to model soil-chemical contaminant interactions. Get a geochemist to help you.
- p.37 Once more, drop acute, and just do chronic health and chronic environment. Keep these indicators separate. I worry about any system that does not translate into a risk essentially computed with standard risk assessment assumptions. (I am particularly worried about spurious risks for non-carcinogens below the RfDs.)

I think you do have to have an algorithm for renormalizing the indicators as the list of chemicals and the information base for the chemicals changes from year to year. Explain what you do to renormalize and give the results with and without renormalization. See my illustrative example in the General Comments section.

- p.41 Having been involved in risk assessment activities on global warming, photochemical smog, stratospheric ozone, and acid rain, I found your willingness to try to expand TRI indicators into these realms a welcome bit of levity at the end of a long, complex report. As your discussion points out, the difficulties are enormous, because of the details of the processes that need to be modeled — and massive efforts have been undertaken by the Agency in modeling risks in these four problem areas. You may wish to coordinate with the people in EPA that are working in these four areas, and from their knowledge, developing suitable indicators (or using the ones they already have developed) as part of a reporting system for strategic environmental planning.

February 28, 1992

Ms. Susan Egan Keane  
Environmental Policy Analyst  
Abt Associates Inc.  
Hampden Square  
Suite 500  
4800 Montgomery Lane  
Bethesda, MD 20814-5341

Dear Susan:

Thank you for the opportunity to review your draft report, "TRI Indicators Methodology," January 9, 1992. I have read the draft carefully, considering the likely uses of the proposed TRI-based indicator and the defensibility of the analytical decisions that you have made.

Before offering my reactions, let me applaud you for both accepting a highly ambitious challenge and making a heroic effort to bring some coherence to very diverse kinds of information and considerations. While I have significant concerns about your proposed approach, they relate more to the search for a single indicator than they do to the specific analytic decisions that you have made. I shall make one general comment, several specific comments, and then offer some local, page-by-page comments.

#### General Comment

My basic concern with the draft is the quest for a single indicator. While the indicator methodology is presented as a risk-based analytical tool, the aggregation necessary in the computation requires numerous debatable assumptions which are outside the province of science. In my opinion, these assumptions are more appropriately made by risk managers, not risk assessors. The draft acknowledges this point to some extent by breaking down the summary indicator into four components: acute human effects, chronic human effects, acute ecological effects and chronic ecological effects.

Since you intend to design a methodology for use on a computer, I recommend that you move toward designing a data base rather than a set of summary numbers. Stated differently, EPA should be in the business of producing an array of summary indicators rather than a single, aggregated indicator. The data base would allow different users (with potentially different interests) to construct their own summary indicators based on their preferred weights and aggregation techniques.

Given the large number of dimensions of environmental quality, EPA is simply asking for trouble if it chooses to produce and report only one summary indicator of environmental quality. By selecting a single indicator that rests on numerous uncertain assumptions, EPA is setting itself up for criticism and embarrassment in the future.

Note that this concern can be addressed by simply recasting the draft report as a project in database design rather than as a project to develop a TRI-based indicator methodology. By

reporting several indicators rather than one summary indicator, EPA would also foster greater public understanding of the multi-dimensional nature of environmental quality.

If such a data base were available, consider how the public might use it. One user might be interested in the total number of people exposed to any chemical thought to be a possible, probable, or known carcinogen. Another user might be interested in the number of people exposed to chemicals that are known human carcinogens. One user might want to select a different baseline than another. Users might differ in what chemicals they want to include or exclude.

It is true that such a data base would be difficult to use, requiring the user to understand the inputs, weights and aggregation options. However, the proposed TRI indicator is easy to use only because it does not require an understanding of its methodology. It is therefore difficult to imagine that its output will be wisely used. Any user who does fully understand the output of the proposed system has all the knowledge necessary to set the parameters in the data base. More importantly, any user who has the necessary knowledge probably would want to construct his or her own summary indicator(s) because users will be faced with different risk management questions and concerns.

### Specific Comments and Questions

1. Your discussion of the current TRI data base should acknowledge that total mass emissions (aggregated over media and sources, for example) is itself a summary indicator based (implicitly) on various assumptions (e.g., each chemical emission is equally bad). By acknowledging that mass emissions is a summary indicator and reporting it next to other summary indicators, insight will be provided into the sensitive weighting and aggregation issues that you have identified.

2. EPA's cancer potency factors, as currently constructed, may be appropriate for screening purposes because it is believed that they provide an upper bound on the true but unknown carcinogenic effect at low levels of exposure. The same factors are not necessarily appropriate for inclusion in an index that purports to represent the status of environmental quality and risk. By combining a bounding estimate on cancer potency with other numbers (e.g., emissions) that are intended to be estimates of real-world quantities, the final scores are not interpretable. It is true that the proposed methodology is intended to capture relative risk and hence the absolute value of the summary indicator is not meaningful. If the cancer potency factor for each chemical were a similar bounding estimate (i.e., each possessed a similar degree [of] conservatism or nonconservatism), then the approach would be valid for purposes of tracking relative risk. Unfortunately, this statement is highly questionable. One of the major themes of *In Search of Safety: Chemicals and Cancer Risk* is that EPA's cancer potency factors are far less conservative for some chemicals than others, even though they are calculated by a somewhat standardized procedure. I urge you to simply acknowledge this point in your discussion, since you have little choice but to use cancer potency factors.

3. The treatment of noncarcinogens is problematic because the approach does not account for the severity of the effect or how the severity (or frequency) of the effect changes with dose. More discussion of this limitation is appropriate.
4. Is it really reasonable to assign the same weight to two chemicals when one has been shown to have several effects and the other only one effect? Your discussion was not very convincing on this point.
5. The draft rejects weighting effects on the basis of their degree of severity as too “subjective.” The draft should acknowledge that an arbitrary, if implicit, assignment of equal weights is also subjective. (Aside: Why not simply let different users set their own weights?) Lurking behind the work group’s decision here might be an implicit bias in favor of hard science (e.g., emissions, toxicity, etc.) relative to soft science (e.g., psychology, economics, and utility assessment). I did appreciate the references to the literature on health status indices and quality of life even though they were not incorporated into the indicator.
6. The draft does not adequately defend its treatment of uncertainty. Why is a factor of 10 appropriate for distinguishing known from possible carcinogens? For a different view of this issue, I recommend that you consult Adam Finkel of Resources for the Future. Why are the exposure uncertainty weights reasonable?
7. Why does the draft focus exclusively on aquatic toxicology?
8. Does the treatment of rural subpopulations effectively give greater per capita weight to sparse populations than for dense populations? Please defend this decision more extensively.
9. Given my (rudimentary) understanding of ecology, I would give big weights to persistence and bioaccumulation. As a consumer of your system, how do I know whether your assumptions are reasonable? Also, shouldn’t the emission weights reflect the baseline health of the particular ecosystem that is about to receive the emission? Some of the chemicals excluded for low toxicity to humans may have significant aquatic toxicity (e.g., copper and zinc). Please double-check this matter. Should you mention non-aquatic organisms or why they are excluded?
10. What is the justification for the specific cutoffs such as 1,000 pounds into each medium and 25,000 pounds of transfers?
11. Shouldn’t the user be allowed to provide extra weight if the exposed population is more than, say, 10% nonwhite? Such data are probably readily available by geographic region. This is a good opportunity to make a statement about the need for consideration of environmental equity.
12. I would discuss section V with Tom McKone of LLL.
13. Need to add a description of what is in the TRI database. Also note some of the problems with the TRI database. For example, is it really believable that cyanide isn’t released at all?

14. Exclusion of non-TSCA chemicals is not a very attractive option. Public is very concerned, for example, about pesticides.

### Page-by-Page Comments

p.1, para 1 replace word “progress” with “health” to convey the sense of a neutral indicator

p.1, para 2 acknowledge difficulty in distinguishing reductions due to regulation and reductions due to other factors

p.2, para 1 (under heading II) note how this paragraph is a powerful argument against toxics use reduction; it is also a nice motivation for your indicator

p.2, para 2 delete third sentence, which is unnecessary

p.2, para 3 the words adequately and adequately enough are used in an authoritative way even though no reasons or evidence is provided in the report to justify these statements

p.4 use default potency values for poorly tested chemicals; check with Alison Taylor at HSPH for a quick-and-dirty method based on acute toxicity tests and/or MTD.

p.5, bottom why not code each chemical as “OTS” and “not OTS” and allow the user to select

p.7, end of first full para — delete “is”

p.8, top of page — sentence beginning “Second,” is true but is not an argument for categories versus continuous values. Please rework the argument.

p.13 acknowledge flaws in IRIS process; contact Kathy Rosica at CMA.

p.13, para 3 how are the potency/toxicity of metabolites handled in the indicator? Am I right that you are focussing [sic] only on parent compounds? Shouldn't this be discussed more extensively

p.14, top — is it 10 percent above the control group or 10 percentage points?

p.14, when you use NOAELs, note how you will select a species for purposes of analysis.

p.15, top RfDs from which route?

p. Figure UpDown typo — actual; typo — portrayed; also make it clear that these are lifetime cancer cases, not annual incidence.

p. Figure CANCERMATRIX            many risk assessors would question whether factor of 100 is appropriate for B1 vs. C. My guess is it should be less than 10. Might check with George Gray at HCRA, Adam Finkel at RFF, and Bob Moolenaar at Dow.

p.17, first full para — how are background exposures handled (i.e., shouldn't weight on threshold pollutant depend upon whether lots of other pollutants with same property are in the area?[]]

p.17, what risk level is used when cancer endpoint is compared to noncancer endpoints? In what fraction of cases does cancer prove to be most sensitive endpoint and how sensitive is this to choice of risk level?

p.17, genotoxicity is arguably less important than cancer

p.18, check with Jack Moore's group at IEHR on your handling of bioconcentration factors

p.18, how about bioaccumulation from soil?

p.20, first full para — expand italicized passage into separate paragraph and provide an example to clarify the point.

p.23, are you talking here about size of current exposed population or size of future exposed population?

p.25, para 1 — the use of word “credibly” is too strong given the limited rationale provided

p.29, note 6 — Have you forgotten effects on fish?

Table SURFACE WATER PARAMETERS — fish consumption varies by region and subpopulation

p.31, para 3, last sentence — would this be illegal disposal?

p.33, top            discussion of 85% is verbose

p.33, para 1 — depth to water?

The “Survey of EPA Scoring and Ranking Efforts” was an excellent piece of work that should be polished and submitted for publication as a separate paper. By placing that piece in the open literature, you will save many people lots of time in the future. The review also provides insight into what is at stake in making progressively more complex scoring systems.

As the work group finishes this effort, I urge both EPA and the work group to devote some time to developing a defense to the “garbage in, garbage out” accusation. I don't think it is a fair charge but it is likely to be made given the large numbers of assumptions and guesses. One way to address it is to design some kind of reality check, which could be used to test the indicator

before it is used as a gold standard. The reality check might be a paper study where a sampling of facilities are subjected to more detailed analysis or the check might involve some real monitoring to provide gross validation of some of the assumptions/predictions. If you move from one to several summary indicators, it is more likely that the “truth” will fall within the range of reported indicators.

Let me conclude by praising the work group for its creative effort to solve a very challenging problem. In my opinion, the most important step you need to take is in the transition from “an indicator” to a “data base” or “data system.”

Thank you very much for the opportunity to comment on your efforts and please don't hesitate to contact me if you desire additional feedback (or encouragement!).

Sincerely,

John D. Graham, Ph.D.  
Director