


United States
Environmental
Protection Agency

Science Advisory
Board (1400)
Washington DC

EPA-SAB-DWC-99-016
August 1999
www.epa.gov/sab

 **AN SAB REPORT ON THE
NATIONAL CENTER FOR
ENVIRONMENTAL
ASSESSMENT'S COMPARATIVE
RISK FRAMEWORK
METHODOLOGY**

**A Review by the Drinking Water
Committee**



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

August 12, 1999

EPA-SAB-DWC-99-016

The Honorable Carol Browner
Administrator
United States Environmental Protection Agency
401 M Street, SW
Washington, DC 20460

Subject: An SAB Report on the National Center for Environmental
Assessment's Comparative Risk Framework Methodology

Dear Ms. Browner:

This Report was developed by the Drinking Water Committee (DWC) of the Science Advisory Board (SAB) in response to interactions with the Agency's National Center for Environmental Assessment (NCEA) during the December 1998 and February 1999 DWC meetings. This Report provides the SAB's reactions to this important effort.

The DWC was pleased that the Agency has taken an important step in beginning the development of an integrated and structured approach for considering complex environmental issues. NCEA illustrated the methodology by evaluating alternative treatment approaches for reducing endemic *Cryptosporidium* risk relative to chemical risks associated with drinking water treatment. We believe that these approaches are needed both from the perspective of science, as well as from the perspective of good environmental public policy making in that they will help to make that process more transparent to the individuals who make up our democratic society.

Even though the Committee is encouraged by the progress on this evaluative approach, it is important to recognize that significant work remains to be done on the methodology. Much of the prior work that has been done is based on narrow evaluations of very specific medical and public health interventions. It is not clear that the approaches for developing common metrics in these settings are appropriate for making environmental health decisions. A specific issue in this regard is the need to insure that the metrics are based on the opinions of an informed public, not just professionals, scientists, or groups that have special interest in the problem. Also, significant problems exist when we bring forth diverse sets of seemingly relevant health risk data in a form that can be aggregated for evaluation

using common metrics. As the DWC's review of the case study demonstrated, significant efforts will need to be focused on the peer review of the individual elements of these analyses (i.e., microbial risks, chemical risks, and the reliability of metrics used), as well as the overall outcome, to ensure that critical information is not lost or misused in such integrative assessment approaches.

The DWC found that the case study provided to illustrate the application of this methodology was very useful for identifying weaknesses in the methodology that are obscured by the somewhat dry textual descriptions of the approach. Despite these concerns we encourage EPA to further explore this methodology with additional case studies in order to help it identify additional areas for improvement. The Committee stands ready to provide additional review and assistance as EPA further develops the basic approach that was outlined by EPA in the documents submitted for this review. We look forward to the response to these comments from the Assistant Administrator for the Office of Research and Development.

Sincerely,

/signed/
Dr. Joan M. Daisey, Chair
Science Advisory Board

/signed/
Dr. Richard J. Bull, Chair
Drinking Water Committee
Science Advisory Board

NOTICE

This report has been written as part of the activities of the Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

Distribution and Availability: This Science Advisory Board report is provided to the EPA Administrator, senior Agency management, appropriate program staff, interested members of the public, and is posted on the SAB website (www.epa.gov/sab). Information on its availability is also provided in the SAB's monthly newsletter (*Happenings at the Science Advisory Board*). Additional copies and further information are available from the SAB staff.

ABSTRACT

The Drinking Water Committee (DWC) of the Science Advisory Board (SAB) reviewed a methodology developed by the US Environmental Protection Agency's (EPA) National Center for Environmental Assessment, Cincinnati (NCEA) entitled *Comparative Risk Framework Methodology and Case Study*. The document presents a methodology intended for analyzing, and describing in comparable terms, disparate health risks associated with alternative drinking water treatment approaches. The Committee supported the continued development of this method and the research necessary to allow its further development.

The Committee noted that the proposed methodology presents a potentially powerful tool that provides a structural framework for identifying important variables that influence the nature and extent of complex environmental problems. The case study that was conducted to illustrate the method's application, while demonstrating its promise, highlighted the difficulties that can be anticipated when such a framework is applied. The Committee suggested that with further development, the Comparative Risk Framework Methodology has the potential to provide valuable insights to officials responsible for local and national decisions on the most appropriate intervention to apply to control human health risks associated with drinking water. The text of the report provides advice that highlights the further efforts that will be necessary for its development and use by the Agency.

Keywords: Comparative risk, economic metrics, common metrics, risk-risk tradeoffs, drinking water risk comparisons

**U.S. Environmental Protection Agency
Science Advisory Board
Drinking Water Committee
Comparative Risk Framework Methodology Panel**

Chair

DR. Richard J. Bull , Battelle Pacific Northwest Laboratories, Richland, WA

Members

Dr. David B. Baker , Water Quality Laboratory, Heidelberg College, Tiffin, OH.

Dr. Mary Davis, Department of Pharmacology & Toxicology, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, WV

Dr. Yvonne Dragan , Ohio State University, Columbus, OH

Dr. John Evans , Program in Environmental Science and Risk Management Harvard School of Public Health, Cambridge, MA.

Dr. Anna Fan-Cheuk , California Environmental Protection Agency, Office of Environmental Health Hazard Assessment, Oakland, CA

Dr. Lee D. McMullen , Des Moines Water Works, Des Moines, IA.

Dr. Christine Moe, Assistant Professor, Department of Epidemiology University of North Carolina, McGavran-Greenberg, Chapel Hill, NC

Dr. Charles O'Melia , Department of Geography and Environmental Engineering, The Johns Hopkins University, Baltimore, MD

Dr. Gary A. Toranzos, Department of Biology, University of Puerto Rico, San Juan, PR

Dr. Rhodes Trussell , Montgomery Watson, Pasadena, CA

Dr. Marylynn V. Yates , Department of Soil & Environmental Sciences, Geology, University of California, Riverside, CA

Consultants

Dr. Judy A. Bean , Children's Hospital Medical Center, Cincinnati, OH.

Dr. Lenore S. Clesceri , Rensselaer Polytechnic Institute, Materials Research Center, Troy, NY

Dr. Richard Gilbert , Battelle Washington, Washington, DC

Dr. Winston Harrington , Resources for the Future, Washington, DC

Dr. Edo D. Pellizzari , Research Triangle Institute, Research Triangle Park, NC

Dr. Joel G. Pounds , Institute of Chemical Toxicology, Wayne State University, Detroit, MI.

Dr. Verne A. Ray , Medical Research Laboratory, Pfizer Inc., Groton, CT

Science Advisory Board Staff

Mr. Thomas O. Miller , Designated Federal Officer, US EPA, Science Advisory Board, 401 M Street, S.W. (1400), Washington, DC

Ms. Dorothy M. Clark, Management Assistant, US EPA, Science Advisory Board, 401 M Street, S.W. (1400), Washington, DC

TABLE OF CONTENTS

1. Executive Summary	1
2. Introduction and Charge	3
2.1. Introduction	3
2.2. Charge	6
3. Comments on the Framework.	9
3.1. General Comments on the Comparative Risk Framework Methodology	9
3.2. Specific Comments on the Comparative Risk Framework Methodology	13
3.2.1. Combining Risk Assessment and Cost-Effectiveness Approaches ..	13
3.2.2. Step 1: Initial Risk Assessment	15
3.2.3. Step 2: Translation of Risk to Health Conditions	16
3.2.4. Step 3: Translation of Conditions to a Common Metric	19
3.2.5. Step 4: Cost Evaluation	25
4. The Case Study	26
4.1. General Issues for the Case Study	26
4.1.1. Is the Case Study an Appropriate Application of the Framework Methodology	26
4.1.2. Statistical Approaches and Uncertainty	28
4.1.3. Case Study Limitations	30
5. Results	31
5.1. Are the Results Presented Appropriate	31
6. Research Needs	32
6.1. Are the Research Needs Clearly Defined	32
6.2. Research Needs to Support Continued Development of the CRFM	33
6.3. Research Specific to the Problem Evaluated	34
7. General Recommendations.	35
7.1 Other Specific Recommendations	35
Appendix A. Specific Comments on the Case Study	A-1
1.Engineering/Water Treatment.	A-1
a. Treatment Scenarios and Cost	A-1
b. Treatment System Efficacy	A-2
2. Risk Characterization	A-4
a. Cancer Risks	A-4
b. Microbiological Risks	A-5
3. Chemical Dose-Response Assessment	A-8
a. Appropriateness of the Discussion of Mechanistic Toxicity for Chemicals	A-8
b. Dose-response Analysis of Individual DBPs	A-9

c.	Appropriateness of Assumptions, Techniques and Models for DBPs	A-9
d.	Usefulness of Mixtures Risk Assessment Approach	A-10
e.	Appropriateness of Assessment of Risks of TOX	A-11
f.	Appropriateness of DBPs Addressed	A-11
g.	Use of Epidemiology Data	A-12
4.	Exposure	A-13
a.	General Issues Related to the Evaluation of Exposure	A-13
b.	Distribution of Drinking Water Consumption Rates	A-14
c.	Unheated Drinking Water Fraction	A-15
d.	Pathogen Survival Relative to Preparation Methods	A-15
e.	DBP Changes in the Distribution System	A-15
f.	Other Routes of Exposure	A-16
5.	Health Conditions	A-16
a.	Health Conditions and Cryptosporidiosis	A-16
b.	Appropriateness of Developmental, Reproductive and Cancer Risks	A-16
c.	Uncertainties in the Process	A-17
6.	Common Health Metric	A-18
a.	QALY Appropriateness	A-18
b.	QALY Assignment	A-18

AN SAB REPORT ON THE NATIONAL CENTER FOR ENVIRONMENTAL ASSESSMENT'S COMPARATIVE RISK FRAMEWORK METHODOLOGY

1. EXECUTIVE SUMMARY

The Drinking Water Committee (DWC) of the Science Advisory Board (SAB) reviewed a methodology developed by the US Environmental Protection Agency's (EPA) National Center for Environmental Assessment, Cincinnati (NCEA) entitled *Comparative Risk Framework Methodology and Case Study*. The document presents information on a Comparative Risk Framework Methodology (hereafter, the CRFM or the methodology), that would be used for analyzing, and describing in comparable terms, disparate health risks associated with alternative drinking water treatment approaches. The Committee appreciates the efforts of the NCEA staff in developing this integrated approach for evaluating complex environmental issues and supports the further development of this methodology and instituting the research necessary to support its application by the Agency.

The proposed methodology is a potentially powerful tool that provides a structural framework for identifying important variables that influence the nature and extent of complex environmental problems. The case study that was conducted to illustrate the application of such a methodology served well to identify the promise of the approach as well as the difficulties that can be anticipated when such a framework is applied. With further development, the Comparative Risk Framework Methodology has the potential to provide valuable insights to officials responsible for local and national decisions on the most appropriate intervention to apply to control human health risks associated with drinking water. This SAB report provides advice on the methodological and data development efforts that will be necessary to further develop this approach for more extensive use by the Agency.

One important need is for the Agency to more clearly and concisely describe how the issue of uncertainty in the model and in data used in the model is accommodated throughout the method. The Agency should provide a complete and detailed description of the methods used to conduct the uncertainty and sensitivity analyses. Specifically, it should state how parameters are determined to be uncertain or variable, the methods used to propagate the uncertainty of the model structure, the methods used to propagate the uncertainty or variability of model parameters in order to approximate the cumulative probability distribution of the model output. An outcome of this analysis should be guidance on how to determine which parameters contribute the most to the overall uncertainty in the model output.

It is extremely important that the Agency abandon the use of upper bounds of risk in these analyses. Central estimates should be used or systematic error will be introduced into the overall risk analysis. This does not preclude the use of confidence intervals in the uncertainty analysis.

Additional important issues to be made a part of the CRFM through methods, guidance, or data development are:

- a) a way to add methods that specifically address the risk of disease outbreaks from drinking water system upsets as well as chronic low-level microbial risk,
- b) additional definition on the processes used to convert continuous data into risk estimates and then to translate risks into health conditions recognizable to the lay public,
- c) research that could establish valuations of health conditions by the public (e.g., Quality Adjusted Life Year–QALY) based on infectious disease morbidity and mortality associated with an environmental risk. There is a need to broaden evaluations in environmental health rather than relying on prior work with medical interventions and preventive medicine.
- d) a better method for dealing with acute nonfatal disease in the derivation of the common metric,
- e) more succinct guidelines for translating adverse health effects from animal experiments to human conditions, especially guidance on how mechanism or mode of action is to be used in these translations.
- f) better foundation for the criteria used to determine incremental exposures from disinfection including disinfection byproduct (DBP) identification, distribution of disinfection byproduct concentrations over time and space, data aggregation for scaling up to the national level, and use of contaminant occurrence information in risk assessment,
- g) a way to consider source water characteristics as a major variable in analyses that go beyond consideration of single water supplies,
- h) ways to consider microbial risks in significantly greater depth. *Cryptosporidium* is not a good prototype for other waterborne pathogens due to its resistance, its behavior in the distribution system, and its minimal risk for secondary spread (also, different classes of microbes have differing sensitivities to different interventions, and different individuals have different susceptibility to disease (genetic susceptibilities and generalized susceptibilities such as those based on age).

Additional specific comments are included in the sections that follow in this report of the Science Advisory Board's Drinking Water Committee.

2. INTRODUCTION AND CHARGE

2.1 Introduction

The Drinking Water Committee (DWC) reviewed the US Environmental Protection Agency (EPA) National Center for Environmental Assessment, Cincinnati (NCEA) document entitled *Comparative Risk Framework Methodology and Case study*. The document presents information on a comparative risk framework methodology (hereafter, the CRFM or the methodology) that would be used for analyzing, and describing in comparable terms, disparate health risks associated with alternative drinking water treatment approaches.

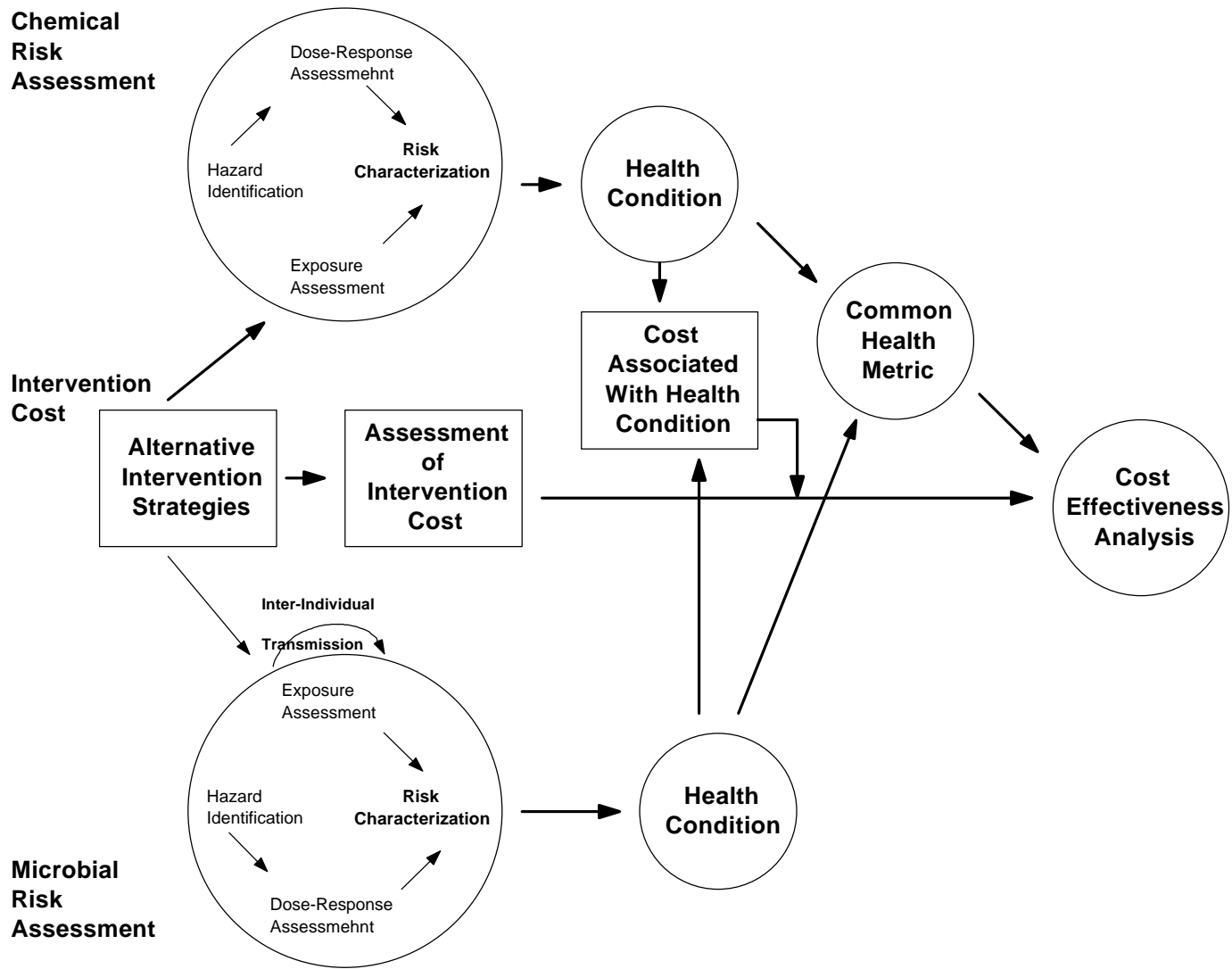
The NCEA made a number of suggestions about what they consider to be the utility and importance of the methodology. They suggested that:

- a) "Use of the CRFM should be beneficial to the local water purveyor who must evaluate treatment options."
- b) "Use of the CRFM should also be beneficial to risk assessors and managers at the national level who must develop data and draft regulations."
- c) "The CRFM...can help provide a sound scientific basis for determining whether or not to go beyond the November 1998 Stage 1 DBP rule to additional regulations for DBPs or microbes."
- d) "EPA believes that the proposed CRFM will assist the Agency in determining the balance between adequate water treatment to control and minimize microbial risk and the creation of unacceptably high levels of countervailing risks from DBPs. The CRFM presented here is intended to support and strengthen traditional and existing risk assessment and risk management activities."

The CRFM integrates cost-effectiveness analysis, as applied to public health interventions, with the 1983 NAS Risk Assessment Paradigm. Figure 1 provides a graphic overview of the basic CRFM. First, traditional risk assessments are conducted for the microbial and DBP contaminants in treated water. Next, to compare these chemical and microbial risks, the effects or consequences described in the risk characterization are translated or expressed in terms of measurable human health conditions, such as cases of cancer, and infection and illness from infectious diseases. The range of potential human health conditions are then converted into a common health metric. In this framework, a Quality Adjusted Life Year (QALY) is used as the common metric to capture changes in the length and quality of life associated with the different health conditions. In the final stage of the framework, alternative strategies are compared by assessing their expected impact on health and economic outcomes.

In addition, the document includes a case study to illustrate how the methodology might be applied to evaluate the health risk tradeoffs associated with two alternative water treatment approaches that focus on reducing risks from the protozoan, *Cryptosporidium parvum*. According to the authors, "The case study is not intended to provide definitive answers

because changes to the specific assumptions used in the case study could alter the results.”
However, it



“...can be

Figure 1. Comparative Risk Assessment Framework Overview (NCEA, 1998).

used to recognize which parameters and assumptions most affect analysis results, while identifying data gaps and uncertainties.”

To consider the potential utility of the CRFM, it is important to understand the statutory process that EPA must follow to regulate drinking water contaminants. First, and foremost, EPA must ensure that drinking water treatment is protective against waterborne microbial diseases. This leads to treatment requirements that control waterborne pathogens in processed water and in the systems used to distribute this drinking water.

EPA is also required to control risks from byproducts of the drinking water treatment process (i.e., disinfection byproducts—DBPs). The EPA Administrator must publish a maximum contaminant limit goal (MCLG) and promulgate a National Primary Drinking Water Regulation (NPDWR) for physical contaminants:

- a) that may have an adverse effect on the health of persons,
- b) that occur in public water systems with a frequency and at levels of public health concern, and
- c) for which regulation presents a meaningful opportunity for health risk reduction.

MCLGs are to be set at a level at which no known or anticipated adverse health effects occur (allowing for an adequate margin of safety). The regulation must specify a Maximum Contaminant Level (MCL) which is set as close to the health goal (MCLG) as is feasible with the use of best technology, treatment techniques, and other efficacious means available while taking cost into consideration. If the determination of contaminant levels is not economically or technologically feasible, EPA can promulgate a Treatment Technique in lieu of an MCL in its regulation. The Treatment Technique must prevent adverse health effects to the extent feasible.

To deal with risk/risk trade-off concerns, the Administrator is permitted to establish an MCL for a contaminant at other than a feasible level if the technology, treatment technique, or other means used to determine a feasible level would result in an increase in health risk by increasing concentrations of other contaminants in drinking water or by interfering with the efficacy of other treatment techniques or processes used to comply with drinking water regulations.

Finally, if a regulation specifies an MCL, EPA must publish and seek comments on analyses that EPA conducts on:

- a) quantifiable and nonquantifiable health risk reduction benefits likely to occur as a result of treatment to comply with the level,
- b) quantifiable and nonquantifiable health risk reduction benefits likely to occur as a result of reductions in co-occurring contaminants associated with the new regulation,
- c) quantifiable and nonquantifiable costs likely to occur as a result of treatment to comply with the level (monitoring, treatment, other),

- d) incremental costs and benefits associated with each alternative MCL considered,
- e) effects of the contaminant on the general population and groups within the general population (infants, children, pregnant women, elderly, individuals with history of serious illness, other subpopulations at greater risk),
- f) any increased health risk as a result of compliance (including risks from co-occurring contaminants), and
- g) other relevant factors (quality/extent of information, uncertainties in the analyses in these special analyses, factors with respect to the degree and nature of the risk).

2.2 Charge

The Agency charge to the Drinking Water Committee asked for comments on both the methodology and the case study. Generally, the Agency asked if:

- a) the purposes of these two components are clearly presented and the document's objectives/ purposes consistent with the rest of the document? Do the materials in Chapters 1 and 2 sufficiently orient readers to the rest of the document? Is the information presented in the document clear?

For the Comparative Risk Framework Methodology section, the Agency asked if:

- a) it is appropriate to combine the NAS 1983 risk assessment paradigm with a cost effectiveness analysis approach? Further, they asked if the method would be useful for environmental and public health decision makers faced with alternative intervention technologies and disparate risks, and if the methodology was presented in a logical format?
- b) initial application of the NAS 1983 risk assessment paradigm to evaluate risks from chemicals and microbes, separately is an appropriate first step, and if other steps should be evaluated?
- c) the second step, translating predicted, separate, risks into human health conditions which are measurable on the population level is appropriate, or if other interim steps should be included?
- d) the third step, translating human health conditions into a common metric, is appropriate or if other interim steps should be included?
- e) the final step, evaluating the financial costs of treatment technologies, medical costs, and a cost effectiveness ratio is appropriate, or if there are other interim steps that should be evaluated?

The Agency prepared the case study to demonstrate the use of the CRFM to evaluate and compare realistic, central tendency risk estimates of both DBPs and one microbial

contaminant, *Cryptosporidium parvum*, across various treatment technologies. EPA asserts that the case study uses hypothetical but plausible assumptions for source water concentrations of *Cryptosporidium*, the efficacy of each treatment system in eliminating this pathogen, concentrations of selected DBPs in finished drinking water for each treatment system, the size of the population served by the facility, the size of the immunocompromised group within the general population, the distribution of drinking water consumption rates, dose-response curves for the DBPs and *C. parvum* exposures, and the financial costs of the treatment options. EPA used a response addition model in the case study because of data limitations on the mechanism of action for the majority of DBPs. Other approaches, for example, another mixtures approach such as dose addition or an assessment of individual chemical constituents of the mixture, could have been developed.

For the case study, the Agency asked if:

- a) this is an appropriate application of this methodology; the initial assumptions are plausible; other factors should be considered within the scope of the case study?
- b) the statistical methods and assumptions used to develop parameter distributions are reasonable and correctly applied; the simulation procedures and sensitivity analyses are appropriate for the data; there are other statistical techniques that could be applied to improve the CRFM?
- c) the limitations of the case study are appropriate and clearly presented? There are other significant limitations not identified?
- d) the choices of water treatment scenarios for evaluation relative to the “real world” scenarios are appropriate? The construction and operation cost estimates adequately reflect real-world facilities and point-of-use device costs?
- e) the accuracy of water treatment system efficacy estimates, including the assumed distribution are accurate and the treatment system assumptions concerning DBP concentrations in finished drinking water are appropriate and reasonable?
- f) overall, the cancer risk estimates, based on exposures to pathogens and DBP mixtures, are plausible and reasonable given the goals and limitations of the case study; the uncertainties presented in the assessments of pathogens and DBP mixtures are adequately described and characterized; there are significant uncertainties that are not identified; the presentations and discussions of variability appropriate?
- g) the assessment of microbial risk based on only a single type of microorganism, *Cryptosporidium parvum*, is appropriate; other organisms should be evaluated with this approach in future applications of the CRFM; the identified hazards related to exposure to the protozoan parasite, and the probabilities of infection and disease conditional on exposure appropriate for the two populations considered (i.e., the general population and the AIDS subpopulation), are appropriate; all possible outcome categories have been identified; the scientific basis of the dose-response model for predicting pathogen risks is well-founded and adequately described and supported; other subpopulations (e.g., the elderly) should be evaluated in future applications of the CRFM?

- h) the discussion of mechanistic toxicity data for individual chemicals and/or mixtures is appropriate?
- i) the assumptions, techniques and models used in the estimation of risks posed by DBP mixtures, including both statistical and biological considerations, are appropriate?
- j) the mixtures risk assessment approach for comparing health risks across drinking water treatment scenarios and levels of DBPs is useful?
- k) the assessment of risks posed by the unidentified halogenated fraction (TOX) is appropriate?
- l) the appropriate DBPs have been addressed and if it is appropriate to not address the carcinogenicity of chloroform in the analysis because of the evidence (Drinking Water NODA, EPA, 1998) indicating that this chemical's mechanism of action exhibits a threshold?
- m) the DWC would offer suggestions for the use of epidemiologic data in future applications of the methodology?
- n) the distribution of drinking water consumption rates are appropriate?
- o) the identification of the fraction of unheated drinking water consumed is appropriate?
- p) the assumption that the pathogen would not survive the pathways from tap to consumption is valid?
- q) the validity and significance of the assumption that DBP concentrations do not change as a result of transport through the distribution system and many pathways through which water is consumed (e.g., boiling to prepare tea)?
- r) other routes of exposure should be included (e.g., dermal and inhalation routes) and whether their contribution would be significant?
- s) the assignment and definition of health conditions are appropriate for the progression of cryptosporidiosis?
- t) the assignments and definition of health conditions for developmental, reproductive and cancer risks are appropriate and if latency periods and reversibility issues are handled reasonably?
- u) the uncertainties in the process are identified?
- v) the use of Quality Adjusted Life Years (QALYs) is appropriate?
- w) the assignment of QALYs given the health conditions are appropriate?
- x) the uncertainties related to the use of QALYs in this application are identified?

y) the results are presented appropriately and the interpretations of the results are appropriate? Additional discussion items should be presented?

z) the research needs are clearly defined?

3. COMMENTS ON THE FRAMEWORK

3.1 General Comments on the Comparative Risk Framework Methodology (CRFM)

The draft Comparative Risk Framework Methodology (CRFM) is among the first methods to systematically compare the risk of microbial diseases with diseases that could occur from exposure to disinfection byproducts (DBPs). Thus, the CRFM has the potential to provide a generally useful context for evaluating decision inputs by individual communities, or for EPA risk managers, by comprehensively framing the impacts of alternative forms of regulatory intervention. When fully developed, the methodology will provide transparency and an integrating structure to the analytical process in support of decision making. It will also provide valuable insights for local and national decision makers who must decide upon alternative interventions that can be applied to deliver safe drinking water. For now, though, significant additional research and development will be necessary for both the methodology and for case specific elements that serve as inputs to the methodology.

Figure 2 shows the CRFM structure as viewed by the Drinking Water Committee. Seen this way, the framework consists of several layers and, like computer software, individual kernels in each layer can be developed and examined independently. The methodology section of the EPA document addresses the uppermost layers, Levels 6 and 7. The case study illustrates how the requirements outlined in Levels 1 through 5 might be fulfilled for a particular situation.

The CRFM is a potentially powerful tool that is well suited to helping identify important variables that are presented by complex environmental problems. The methodology can be a significant advance for evaluating environmental problems and the Committee is enthusiastic about the overall analytical structure that the framework provides, in particular, the concepts described in Levels 6 and 7 (i.e., the description of the "quality-adjusted life year"--QALY method). For these levels, major remaining tasks are to compare alternative choices for the common health metric to determine which is most suitable, and then to develop defensible weighting parameters for the metric once selected. The description of the QALY method for comparing health risks is clear and the arguments for the use of some common measure of health impact reflecting both longevity and quality of life are persuasive. Even though the Committee is not questioning whether the QALY is the most suitable metric, it believes that alternatives should be considered, adequately discussed in the document, and a clear case made for the metric selected.

The case study served well to identify both the promise of the CRFM and the complexities that can be anticipated in its application. However, before the method is adopted for the analysis of national policy issues (such as the microbial risk/disinfection byproduct risk issue) further methodological development will be necessary to ensure that the complexity

introduced by site-to-site variation in raw water quality, treatment costs, and community values can be accommodated. Refinement of Levels 6 and 7, alone, will not be sufficient for successful application. Data and methodology gaps will also need to be filled at Levels 1 through 5. Fortunately the bulk of the nation's current DBP/Disinfection research is directed at gathering information that will support the development of methodologies at levels 1 through 4.

The CRFM uses a common metric to compare all potential risks associated with situations being evaluated. However, there is a need for research to develop measures that are valid for environmental health issues. Methods to translate the risks estimated in Level 4 into the probability distribution function (PDF) of health outcomes required in Level 5 will be critical (especially the identification of relevant health outcomes in human populations on the basis of data from animal studies). Another key issue needing further attention is the treatment of uncertainty -- particularly model uncertainty.

Many CRFM applications will require analysts to predict human disease outcomes based on evidence from animal toxicology studies. Because the "policy importance" of a disease in the QALY approach is determined, in part, by the loss of quality adjusted life expectancy associated with the disease, the correct identification of disease conditions will be important. However, the relationship between the effects seen in animal studies and the diseases apparently detected in human populations exposed to DBPs is not always clear. For example, the case study assumed that the diseases resulting from human exposure to disinfection byproducts would be bladder and rectal cancer -- i.e., the diseases which have been seen in human epidemiological studies. However, the quantitative risk estimates for these diseases were based on extrapolation of results from animal studies which, instead of observing these disease outcomes, observed liver and kidney cancers in study animals. It is not clear that these outcomes from animal studies will adequately extrapolate and translate to the equivalent QALY weights in humans.

The Committee understands and agrees with the logic underlying the decision to compute the present value of health impacts. Further, we believe that in the current case study the particular choice of a discounting rate would not substantially influence the results. However, in other cases the choice of the discounting rate is likely to be important. Because of this decision, and in view of both the wide-spread misunderstanding of the rationale for discounting health impacts, the legitimate ethical arguments surrounding this issue, and the political sensitivity of the issue, we recommend that the document include a more prominent and nuanced discussion of the general issue and that it provide stronger support for the particular choice of 3% as the discount rate used in the analysis underlying the case study.

The Agency document does not describe the methods used to conduct the uncertainty and sensitivity analyses in sufficient detail. A complete and detailed description should be provided that discusses how parameters are determined to be uncertain or variable, the methods used to propagate the uncertainty of the model structure, the methods used to propagate the uncertainty or variability of model parameters in order to approximate the cumulative probability distribution of the model output, and the methods used to determine which parameters contribute the most to the overall uncertainty in the model output. The Agency may wish to place the detailed description in an appendix. However, the major steps of the process should be explained in the body of the report with the aid of flow-charts or similar graphical methods.

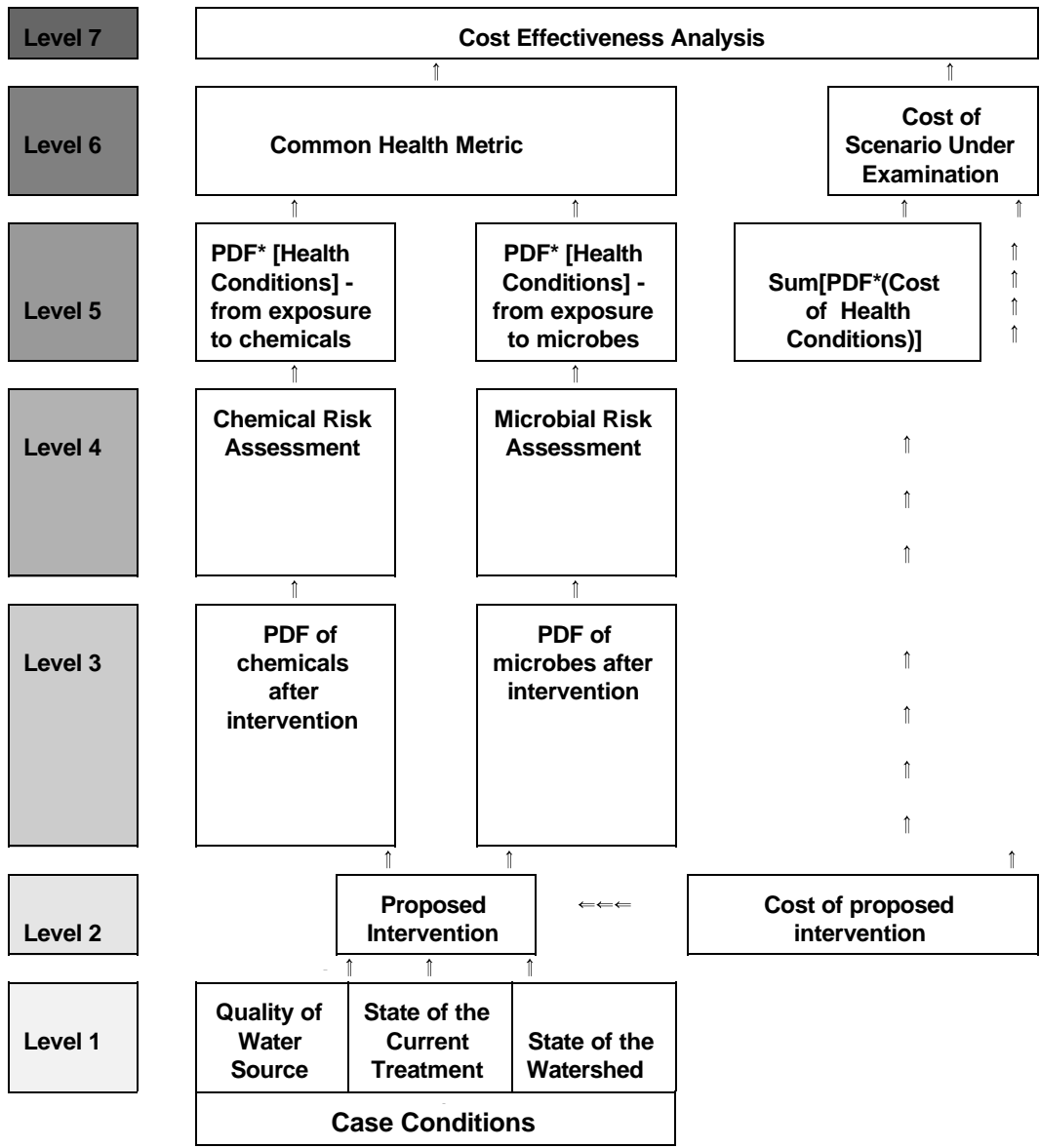


Figure 2. DWC Overview of the Proposed EPA Comparative Risk Framework Methodology

*PDF = Probability Distribution Function

Of special concern to the DWC is the lack of an adequate discussion, in both the general methodology and the case study, of how to assess and incorporate the magnitude of model uncertainty in the total uncertainty of the model outputs. The case study focuses on propagating uncertainty in some model parameters, but it appears that the model structure itself is assumed to be known with certainty. This assumption is not realistic and some method is needed for including model uncertainty in the uncertainty analysis. Also, additional effort is needed to separate and properly take into account the effects of variability across populations and uncertainty in models and parameters. A general methodology for this process should be developed and illustrated in the case study. Other issues that should be addressed in more detail are the rationale for the selected probability distributions for uncertain parameter values and variable quantities, the methods used to take into account the correlations among key parameters of the main model and any submodels, the methods that should be used to obtain expert opinion about alternative model structures, exposure scenarios, and the probability distributions of uncertain or variable quantities, and when methods such as Latin Hypercube Sampling for generating random numbers from specified probability distributions of variable or uncertainty quantities are preferred to simple random sampling. Just as important is that parameters involved in assigning utility to changes in health status, such as QALY weights, may be uncertain and should be considered as such in the uncertainty/sensitivity analysis.

The Committee suggests that the methodology be cast in the most general case. Effects and exposure data vary and they generally follow a distribution rather than being dichotomized. For example, sensitive populations should be cast as a distribution of sensitivities as the baseline in the methodology. As the CRFM is applied to specific cases, it may be necessary to dichotomize the risks. For example, the dichotomization of human sensitivity to *Cryptosporidium* into "normal" versus HIV-infected individuals in the case study was appropriate because that was the nature of the data that were available (i.e., a comparison of a specific population relative to the norm). But it must be remembered that HIV-infected individuals make up only a portion of the sensitive population. This had implications in the case study because treatment of sensitivity only in this selected population tends to minimize the effectiveness of better water treatment.

The methodology does not effectively address the problems of outbreaks of microbial disease. Outbreaks are more visible to the public at large, and their effects differ in that they are not just estimates, but are confirmed reality. Therefore, the methodology should consider the risks from outbreaks of waterborne disease separately from the risks of endemic waterborne disease. In general, endemic waterborne disease may be more related to distribution system problems while outbreaks of waterborne disease are usually due to treatment failures. Data from Payment's studies (1991, 1997) suggest that viruses may be responsible for a large portion of endemic waterborne disease. Outbreaks of waterborne disease in the US are due to both protozoan and viral agents. Most identified etiologic agents of waterborne disease outbreaks are protozoa (*Giardia* and *Cryptosporidium*). However, the etiologic agent is never identified for approximately half of the waterborne disease outbreaks. These outbreaks are likely due to viral agents based on their epidemiologic characteristics. They likely are not identified because of our limited capacity to detect viral agents in clinical and environmental samples.

It is the Committee's opinion that omitting outbreaks will have real implications for the risk manager. The measures that will most effectively address endemic disease are not

necessarily the measures that will most effectively address outbreaks. Reducing endemic disease requires lowering the mean level of pathogens in the water supply over long periods of time. Preventing outbreaks requires improving the reliability of the barriers between the sources of pathogens and the consumer's tap.

The Comparative Risk Framework Methodology developed by the Agency is a broad outline of an approach to compare risks of a disparate nature. However, application of the method in an illustrative case study revealed that there are additional layers of infrastructure needed that require careful consideration before the method can be successfully implemented. The Committee recommends that EPA conduct additional case studies of substantially different conditions than those reflected in the existing case study, in order to ensure that solutions are developed for the practical implementation problems noted by the DWC. Properly chosen, these case studies might provide useful insights for decision-making. In this effort, the DWC recommends that the elements of the analyses (e.g., DBP risk, microbial risk, and economics) be subject to rigorous peer review. This would ensure that all pertinent information is captured and that the uncertainties are appropriately propagated in the overall application of the methodology.

3.2 Specific Comments on the Comparative Risk Framework Methodology

The subsections that follow address specific EPA charge questions on the methodology. In Section 4 of the report, we provide general comments on the case study. Later, in Appendix A, we provide comments on details of the case study that are not necessary to address in the body of the document, which addresses the Comparative Risk Framework Methodology itself.

3.2.1 Combining Risk Assessment and Cost-Effectiveness Approaches :
The proposed CRFM in Chapter 4 combines the NAS 1983 risk assessment paradigm with a cost effectiveness analysis approach. Is this combination appropriate? Will the CRFM be useful for environmental and public health decision makers faced with alternative intervention technologies and disparate risks? Is the CRFM presented in a logical format? If not, please suggest an alternative organization.

a) General:

The Drinking Water Committee concluded that the combination of risk assessment and a cost effectiveness analysis is appropriate in the methodology, and that it has the potential to be quite useful in giving new insights to decision makers who are considering appropriate types of environmental control, as well as to those prioritizing research efforts that target critical decision making information.

The Committee is aware that in many real cases the framework will fail to strongly differentiate among the alternative control strategies (i.e., on the basis of their cost-effectiveness). If good science and regulatory analysis cannot clearly differentiate among strategies showing one to be superior to another on the basis of economic efficiency, then the decision maker may have a legitimate basis for making the decision on other grounds. It was

apparent to the Committee that authorizing legislation may preclude the application of the CRFM in some situations. For example, it is inappropriate under the Safe Drinking Water Act (SDWA or the Act) to establish maximum contaminant level goals (MCLGs) for a contaminant in this way. The MCLG is to be established strictly on the available health effects information, based on agreed upon policies. On the other hand, the method would be suited to identify issues that are important to establishing a Maximum Contaminant Level (MCL), since the Act allows for consideration of a variety of practical factors in setting the MCL. An example would be a lack of an available (affordable) treatment process to meet the MCLG. The methodology would also have an application in determining the alternatives that would be pursued in obtaining variances or exemptions from MCLs.

The Committee endorses comprehensive frameworks that integrate risk analysis within decision making, and that consider simultaneously the health consequences and the control costs of decisions. However, this endorsement does not imply that we believe cost-effectiveness analysis to be superior to other approaches such as benefit-cost analysis. In fact, we believe that increased communication among the EPA offices that have used or are considering comprehensive frameworks would be beneficial. This might encourage cross-fertilization of ideas and consistency in the approaches taken to evaluate regulatory programs.

b) Specific Comments:

EPA's Office of Policy (OP, formerly the Office of Policy, Planning, and Evaluation) has a long history of leadership in developing and applying a comprehensive economic framework to analyze public decisions—specifically, they have worked to develop benefit-cost methods for evaluating proposed environmental regulations. Conceptually, that method values an improvement to an individual's health status as the maximum sum of money the individual is willing to pay for the improvement.

The CRFM differs from the Office of Policy's typical regulatory analyses because it relies on cost-effectiveness analysis rather than benefit-cost analysis. Only human health effects are considered, and no attempt is made to monetize those effects. Instead, all health effects are valued using a common health metric. Each of the two approaches has strengths and weaknesses and the use of QALYs does not preclude the use of benefit-cost analysis. A discussion among representatives from NCEA, OP, and potential users of the method could be helpful in establishing the place for each approach.

The innovative aspect of this proposal is not on this risk assessment approach *per se*, but in putting diverse information on risk into a framework better suited for decision-making. Essentially, this involves developing a common metric – Quality Adjusted Life Year (QALY—discussed later). While this does not change the fundamentals of the NAS 1983 risk assessment paradigm, one must be sure that the process is internally consistent with arriving at outputs that are compatible with the development of the common metric.

The CRFM team has made a commendable effort to integrate and coordinate the diverse concepts, processes, and data presented in this document. However, the document prepared by EPA was very complex, and the Committee found that its organization and style of presentation made it difficult to easily identify information on specific topics or the mechanics of addressing and folding the components into a comparative risk assessment. An

effective approach might be to organize the general methodology and the case study so that the section and topic headings correspond to the schematic illustrations of the method. It would also be helpful if subsequent versions of the document make more effective use of tables and graphics to highlight critical aspects of the process.

Finally, the Committee emphasizes its concern that the components of these analyses be subjected to appropriate peer review. The complexity of this approach as it has been applied to the microbial and disinfectant byproduct example will be a powerful incentive to ignore data not clearly understood by the analyst. Making certain that the technical underpinnings of the analysis are correct and up-to-date will be the major difficulty in bringing a completed analysis to the decision maker.

3.2.2 Step 1: Initial Risk Assessment

The proposed CRFM (Chapter 4) initially applies the NAS 1983 risk assessment paradigm to evaluate risks from D/DBP chemicals and microbes, separately. Is this an appropriate initial step in the methodology? Are there other steps that should be evaluated?

Risks from disinfectants/disinfection byproducts and microbial agents should be independently evaluated, as should different endpoints such as cancer and reproductive toxicity, before proceeding with a comparative assessment. The 1983 NAS risk assessment paradigm is an obvious place to start for both types of agents. However, there are distinct issues with each type of agent and the degree of understanding of interdependent parameters of risk and disease that must be kept segregated to make the analysis transparent.

A microbial risk assessment paradigm that is consistent with the general 1983 NAS risk paradigm, yet which expands on and shows the unique characteristics of microbial risk can be found in ILSI (1996). Further, a paper by Sobsey, Dufour, *et al.* (1993) presents a systematic strategy for identifying, analyzing, quantifying and characterizing microbial risks. These papers are relevant to the development of the current CRFM approach, and the Committee recommends that the EPA staff also consider them in assessing microbial risks.

There have been many advances in our understanding of how disease is produced by microbial and chemical agents that were taken into account, at least in part, in the document. However, some of the fundamental issues that underlie the use of research data to arrive at estimates of risk are not readily apparent in the document. This is due, in part, to the document's style. The Committee is concerned that some users and reviewers of an analysis based on the framework may not recognize the nature of the assumptions that underlie most risk assessments. Most analysts are unlikely to be sufficiently well grounded in toxicology or microbiology to understand the actual magnitude of some of the uncertainties that can be glossed over in such an analysis. While somewhat mundane and certainly covered in other EPA documents, the Committee considers it to be important to make these technical issues explicit in the CRFM document.

The methodology should also consider the role of secondary person-to-person transmission of waterborne infectious agents in the calculations of disease impact. For many microbial pathogens, secondary transmission is responsible for significant morbidity. Data on

probabilities of secondary transmission can be collected from reports of outbreak investigations and from investigations conducted by the Centers for Disease Control (CDC).

Host susceptibility to infection from various waterborne pathogens is a complex issue and can vary widely depending on the pathogens under consideration. For some pathogens, such as Hepatitis A virus, infection results in life-long immunity. For other pathogens, such as Norwalk virus, the majority of the population has serum antibodies; however, these are not protective and re-infection occurs frequently. For pathogens such as *Cryptosporidium* there is great uncertainty about what method to use to evaluate seropositivity, what constitutes a positive serologic test, what percentage of the population has *Cryptosporidium* antibodies and how the presence of antibodies affects the risk of re-infection, disease occurrence, and disease severity and duration. The Committee recommends that where the uncertainty is large, then the methodology consider a range of host susceptibility in the model. This recommendation also applies to the reinfection period used in the model. When there are no data from human challenge studies or outbreak investigations, it seems prudent to treat the infection period as an uncertain parameter with values ranging from perhaps 3 months to 1 year.

The availability and quality of dose-response data for waterborne pathogens varies. The Committee recommends that, when available, the median infectious dose be used in the CRFM models, and that the uncertainty associated with the estimate be recognized. The median infectious dose can vary significantly, depending upon the strain of the microorganism under consideration.

It is important to recognize the limited availability of important data on pathogen risk. This is especially true in the case of pathogen occurrence. This situation is further exacerbated by the poor sensitivity, recovery, and detection limits of the methods that have been used to develop the occurrence data that do exist. In contrast, DBP occurrence data is relatively large and it is certainly accessible. Thus the methodology may generate a biased result.

3.2.3 Step 2: Translation of Risk to Health Conditions

The second step is to translate the predicted, separate, risks into human health conditions which are measurable on the population level. Is this an appropriate next step in the methodology? Are there other interim steps that should be included?

The translation of the predicted risks into human health conditions is a logical and appropriate next step in the methodology. However, at the February 1999 Committee meeting it became apparent that the methodology did not identify all the assumptions and steps necessary to translate the available data from animal toxicology research into human health conditions. Examples are given below to illustrate these problems. One involves the use of information derived from studies in which parameters were measured as continuous responses in the derivation of risk estimates expressed as stochastic variables, another the use of alternative assessment approaches, and the third the effective use of both animal toxicology and human epidemiology data in the identification of health conditions.

The first problem arises because much of the most useful toxicological research measures continuous responses associated with dosing by some chemical agent (i.e., increasing dose to an individual increases the magnitude of a response in that individual). The Agency frequently uses this type of information in developing MCLs, and NCEA did include such data in this analysis. However, risk is expressed stochastically (i.e., increases in dose increase the number of individuals with the disease, in other words, individuals either respond or they do not respond at a given dose). The conversion between the data types is generally based on some arbitrary assignment of a point on the continuous scale that will be considered the dividing line between positive and negative stochastic responses. This can potentially confuse the derivation of the common metric which ultimately depends upon human evaluations of the quality of life associated with the health conditions ascribed to that positive response.

Therefore, the factors considered in the conversion of continuous to stochastic responses can not be arbitrary. An example of how this has been approached for neurotoxicity endpoints can be found in Kodell et al. (1995). The probability that a given alteration in the continuous variable will lead to the development of an adverse impact that is defined in stochastic terms must be explicitly considered. Then comes the assessment of risk (i.e., the frequency of the adverse impact in a population arrayed against dose or exposure). Finally, the adverse impact must be translated into a disease condition that a lay person can recognize and assign some negative value to which can then be converted into the metric that will be used to normalize different health effects. All these intermediate steps must be explicitly considered for the application of the common metric to be scientifically sound.

Another way of looking at this issue is illustrated by the diagram in Figure 3. Continuous variables in toxicological experiments are measurements of effects on enzyme rates or other protein function that are involved in normal biological functions. These parameters can often be measured within isolated cells or intact animals, but generally not in humans. To be most useful the parameters measured should be involved in the mechanism by which the chemical produces harm. However, there are many factors that affect the expression of such effects as a frank disease. Most important is that the shape of the dose-response curve describing the probability of disease is determined by the distribution of sensitivities of individuals in the population to the modifications in the biochemical system measured. In most cases, the distribution of these susceptibilities in the population will be unknown. As pointed out in the CRFM document, continuous variables can be converted to a population-based figure by some relatively simple mathematical conventions, so something like an incidence term can be developed. However, such arbitrary treatments can trivialize the translation processes because they can gloss over fundamental flaws in our understanding of the mechanisms leading to environmentally induced disease. As such they can introduce substantial uncertainties into the risk assessment process that need to be accounted for by the methodology.

As results of animal studies are integrated with epidemiological studies, it becomes important to understand how the analysis will be calibrated in terms of attributable risk. This will be difficult, but it must always be attempted. Are the projections of risk reasonable considering other recognized causes of a particular disease? Are the projections of the animal data consistent with the epidemiological literature and are the epidemiological data consistent with the toxicological data? The handling of these problems in the CRFM was cursory and not

very defensible.

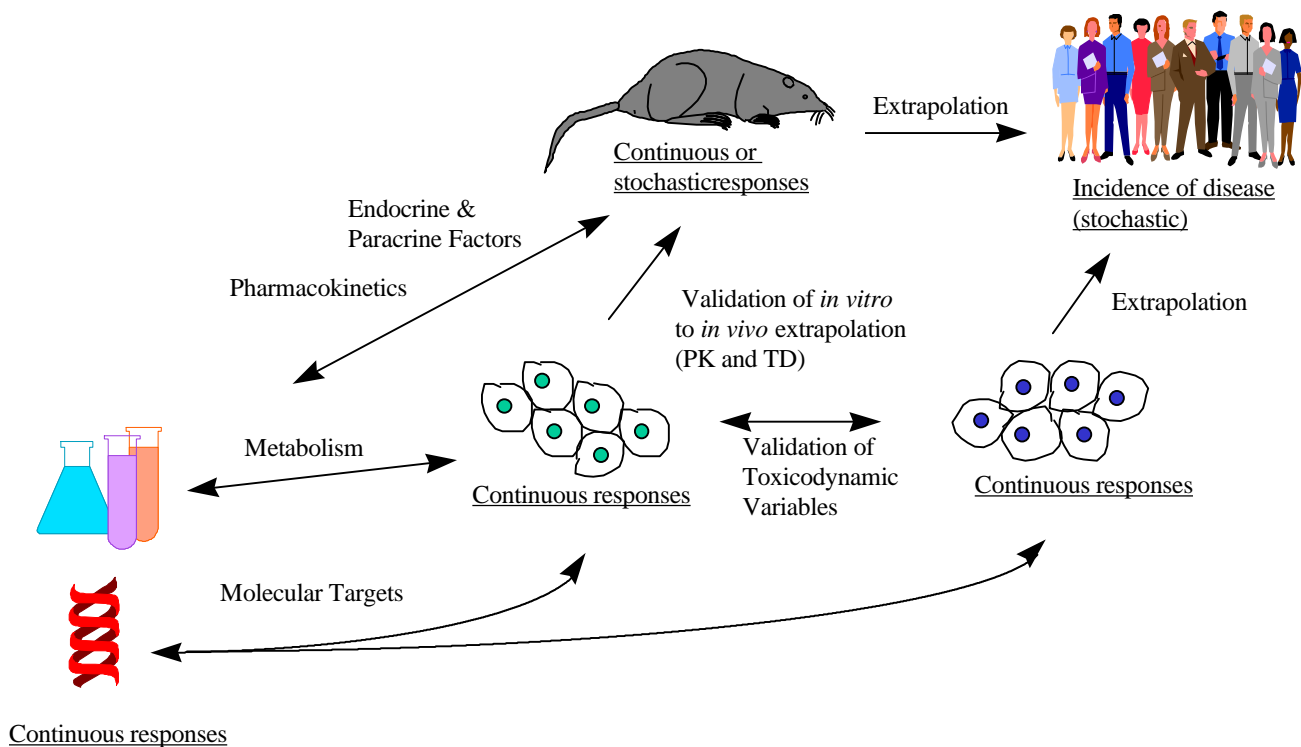


Figure 3. Relationship between different indicators of adverse effects of chemical and physical agents that can be measured. Toxicological research frequently involves measurement of biochemical or molecular parameters whose quantitative relationship to the incidence or severity of disease is not explicitly understood. Measures of these ‘key events’ indicate a genuine potential for producing disease. Such measures are frequently used to develop regulations for individual chemicals. However, if they cannot be converted into a probability that a certain condition will occur in a population, it will be difficult to normalize the effects of agents that produce diverse outcomes.

A second point is that more attention should also be paid to how policy affects the way in which agents are considered in the CFRM. Clearly, the risk manager may be constrained by Agency policy. The Committee encourages the comparison of results from alternative approaches to risk assessment. This would make apparent to the Agency whether certain policy prescriptions would have significant impacts. This might also serve to identify areas where Agency research might be targeted.

The third problem involves translations of a different sort. Animal data were used to estimate quantitative risks, but the results of epidemiological studies were used to develop the common metric (e.g. the Quality of Adjusted Life Years). Part of the problem arises from the very different quantitative risks that are projected from epidemiology studies (Morris et al., 1992; Poole, 1994, 1997) versus those risks based on available toxicological data on DBPs (Bull and Kopfler, 1991). Toxicology provides the type of information that is needed to deal with individual byproducts, so it is possible to estimate risks from the concentration of individual DBPs and proceed accordingly. However, the toxicological data address only a few of the DBPs that are found in disinfected water. Thus, it is entirely possible that both estimates are accurate within their areas of investigation. It is very possible, or even likely, that the two approaches are measuring the effects of different byproducts or different byproduct combinations.

These examples point out the temptation (or even need) to oversimplify that will always be encountered in general cases of comparable complexity. It is imperative that some instructions be provided to indicate the need to comprehensively develop such cases and then to explicitly identify the simplifications that have been made and to include this in the uncertainty analysis. Therefore, it is critical to explicitly address these issues when it is necessary to extend a large body of information on health effects of a number of agents to make predictions about a larger set of unknowns.

The Committee recommends that more explicit guidance be provided in Section 4.4 of the Agency document about decision criteria and paths to be taken in exercising this aspect of the CFRM. The interim steps that are taken to make translations between different data types (e.g., continuous responses underlying “key events”) have to be viewed carefully in terms of their potential impact on the translation of potential human health effects into common measures. In turn, the contribution of the assumptions that these translations have on the uncertainties that are embedded in the final analysis must be acknowledged.

3.2.4 Step 3: Translation of Conditions to a Common Metric

The third step is the translation of the human health conditions into a common metric. Is this an appropriate next step in the methodology? Are there other interim steps that should be included?

The third step, the translation of the human health conditions into a common metric, is a necessary and appropriate step in this cost-effectiveness analysis. The quality adjusted life year (QALY) was the metric chosen by NCEA. Theoretically, this metric would allow quantitative comparisons of different kinds of adverse health effects by quantifying subjective responses to hypothetical questions (an example of the calculation of the QALY is given later in this section).

The QALY concept has been used by international agencies (such as the World Health Organization and UNICEF) to weigh the benefits of various public health interventions in developing countries. For example, in efforts to improve children's health (i.e., reduce childhood morbidity and mortality due to a variety of infectious diseases) such agencies must determine which of several possible interventions (e.g., immunization programs to prevent measles and typhoid, oral rehydration solutions to reduce diarrhea mortality, pesticide impregnated bed nets to reduce malaria, improved water and sanitation to reduce diarrheal diseases, and primary health care clinics) to pursue. Faced with this problem, such agencies have relied on the QALY (or other similar metrics). Because this approach is new for EPA, the SAB strongly recommends that EPA consult with experts from other public health/development agencies who are experienced in this type of analysis.

The CRFM is not the first EPA effort to use a comprehensive economic framework to analyze public decisions having a health component. The EPA's Office of Policy referred to earlier in this document has developed methods to estimate the value of the beneficial effects of environmental programs. A collection of EPA case studies on the use of benefit-cost analysis is contained in Morgenstern (1997). Especially noteworthy are the cases dealing with regulation of atmospheric lead and lead in drinking water. The benefit-cost analysis (BCA) method uses monetary units as the common metric for valuing all benefits, including the health benefits. Conceptually, a given improvement to an individual's health status is valued by the maximum sum of money the individual is willing to pay for the improvement – i.e., the individual is indifferent between having the lower health status with the money and the higher health status without the money. To be clear, the QALY cost of a health condition is “the life years in perfect health” that an individual is willing to sacrifice in exchange for elimination of the health condition under consideration. For example, a 25 year-old who is indifferent between living 50 years with her current ailments and living 47.5 years in a state of ideal health would have a cost associated with her current condition of 2.5 QALYs ($50.0 - 47.5 = 2.5$) and each year in her current condition would be valued at $47.5/50 = 0.95$ QALY/year.

Like willingness to pay (WTP), the QALY is a utility-based measure of health effects. But instead of measuring an individual's indifference between various states of health and money, the QALY measures the individual's indifference between years with various diseases or disabilities and years of “perfect” health. It combines impacts on longevity with morbidity outcomes using an intuitive framework -- i.e., the number of years of life loss that one would trade for a year with any specific ailment.

Each of the two approaches (i.e., cost-effectiveness analysis and benefit-cost analysis) has strengths and weaknesses, and a dialogue between users of each could be profitable to both. The greatest advantage of the dollar metric is its total comparability with costs. Money is the only metric that allows measurement of all the outcomes of a policy intervention in the same metric and thus allows the calculation of “net benefits.” Its biggest disadvantage is that it seems to reduce everything to a matter of dollars and cents. Applications of benefit-cost analysis to health issues frequently run into objections, from lay persons as well as health professionals who are uncomfortable about attaching dollar values to changes in health status.

Some may find cost-effectiveness analysis more “palatable” than benefit-cost analysis because cost-effectiveness analysis does not require that health impacts be evaluated in

monetary terms. Public health decision makers faced with choices among alternative intervention technologies may find it easier to communicate the logic behind their decisions to stakeholders and the public if they rely on cost-effectiveness rather than benefit-cost analysis.

In fact, QALYs might offer a number of other advantages over the direct monetization of health benefits to practitioners of BCA. First, QALYs have undergone extensive study by public health researchers, medical decision analysts and health economists, and there is a large literature to draw on for determining the QALY scores of various health states. In principle, QALYs also have some transferability from one situation to another. QALY weights in one study can in many cases be used with some plausibility in another, avoiding the expense of an additional survey. Finally, the QALY concept offers a way of treating mortality and morbidity consistently within a common framework. This is missing in the usual approach to health benefit estimation, where individual disease conditions are treated separately and on an *ad hoc* basis. In particular, mortality is frequently valued by means of the “value of statistical life,” which is based on individual willingness to pay for very small changes in the risk of dying. QALYs may be a useful device for permitting the comprehensive treatment of mortality and chronic disease in BCA.

Despite these advantages, there are circumstances where cost-effectiveness analysis based on QALYs may be inadequate. For example, where the benefits from regulatory interventions include appreciable impacts on species other than humans or on the environment itself, it becomes necessary to rely on “common metrics” which are more broadly based – such as dollars.

Furthermore, although the QALY is a comprehensive health measure, it has some weaknesses, and it may be weakest in the valuation of acute health effects. The approach taken in the CRFM to estimate QALYs for relatively minor, short-duration illnesses involves many un-tested assumptions. Further research is needed to directly evaluate individual tradeoffs between episodes of short-term illness (e.g., vomiting and diarrhea) and small reductions in life span or small increases in the risk of immediate death. Quite possibly, it is an easier conceptual task to trade off minor acute illnesses against money than against later chronic disease or risk of death.

QALY estimates should not become the sole basis for weighting the strengths and limitations of various water treatment interventions. For example, it may be tempting for some decision-makers to look at the bottom line of the cost-effectiveness analysis and choose the intervention that gives the lowest cost per QALY. Such a choice may provide the most cost-effective health benefit but may miss other important considerations (such as time, or access to/availability of health care, etc.) that need to be weighed in the decision. For example: In 1979, an important analysis by Walsh and Warren concluded that a primary health care package was a much more cost-effective intervention to reduce childhood mortality in developing countries than improved water and sanitation and this analysis formed the basis for WHO policy for several years. However, Okun pointed out that this analysis only looked at childhood mortality and ignored the benefits of improved water and sanitation for free time for women (the main water collectors), improved economic development, improved community organization, improved domestic agriculture, etc.

When EPA uses this approach to weigh the value of several water treatment options,

the agency should explicitly list all possible strengths and limitations associated with the proposed interventions that should be considered in addition to the QALY estimates. In the interventions considered in the case study, home filters may involve aspects of maintenance time and effort and difficulty monitoring finished water quality that are not reflected in the financial cost of the intervention. Water taste is another consideration that is hard to quantify in this type of analysis.

It is important to note that the use of QALYs does not preclude the use of benefit-cost analysis. There is no reason in principle why a dollar value cannot be attached to a QALY. It would represent the individual's valuation of a year of perfect health. It might even be useful to compare the direct valuations of health states in monetary terms with their valuations as determined in a two-step process, first determining their QALY score and then converting to monetary units by means of an estimate of the monetary value of a QALY.

The Committee identified a number of specific issues associated with the calculation of QALYs that need further consideration. Some of these issues simply require more explanation in the text; others may require a more extensive survey of the health cost literature.

a) The elicitation method used to derive QALY weights deserves more discussion. In chapter 4 the Agency poses the time tradeoff as *the* way to calculate QALYs. The CRFM report mentions two alternative approaches, the standard gamble and the use of subjective rating scales, but argued that the flaws in each of these precluded their use.

The time tradeoff method is not accepted as the best method by all researchers. The time tradeoff method seems to embody some hidden assumptions about time preference, considering that the years of perfect health that one trades off necessarily occur at the end of one's expected life. When this question was raised at the DWC meeting on December 10, 1998, the Agency responded that the question of time preference is never brought up during surveys to estimate QALYs by the time-tradeoff method. Is it correct, i.e., are there no debriefing exercises among survey respondents to determine whether they were taking discounting into account in valuing the reduced life span?

b) Which populations should be used to derive QALY estimates of disease severity? Should QALY weights be determined from a random sample of the population; a sample of those suffering from the disease of interest; or a group of experts such as doctors and nurses who treat patients suffering with the disease of interest? All have been used in studies designed to elicit QALY values.

For example, in a study by Hamilton, a source of QALY weights for the CRFM case study, the respondents are parents of children identified as attending publicly funded schools. While one of the three disease outcomes of interest in the CRFM case study was change in fertility, the Hamilton study population, excluded a potentially relevant group of respondents -- i.e., childless adults. This excluded (or at least under-represented) group may value fertility impacts quite differently than other members of the community.

It seems that each approach would have certain advantages and disadvantages. The SAB urges the Agency to more fully consider these and other alternatives and to explain to potential users of the methodology the advantages and disadvantages of each approach.

c) The concept of “perfect health” seems to suffer from a lack of definition. Is “perfect health” different for a 20 year old than it is for a 60, 70, or 80 year old person? How is that taken into account?

d) The report should discuss whether the studies that have been used in the past to establish QALY weights are appropriate for the kind of environmental decision-making envisioned by EPA for the CRFM. QALY weights are established by means of surveys that ask respondents to make tradeoffs about their own health. Would the answers be the same if the respondents were asked to make decisions about public health interventions, in a format that more nearly matched the manner in which these questions are raised in the political arena? This would require questions that, for example, asked respondents whether they would support a program that would prevent X cases of disease A or an equally costly program that prevented Y cases of disease B. Are the authors aware of any QALY studies that have tried to rank interventions in this way, and if so, compared the rankings thus generated with those generated by a set of QALY weights?

e) The report and its authors should be prepared for questions to be raised about the appropriateness and possibly even the legality of using QALYs. Several years ago the State of Oregon proposed to use QALYs to determine the kinds of health interventions that would be reimbursed under Medicaid (i.e., possible treatments were ranked by a value derived by a measure of the improvements to QALYs they achieved divided by the cost of the treatment). The Oregon approach was twice rejected by the US Department of Health and Human Services (DHHS) based on their determination that the Oregon procedure for deriving QALY weights violated the Americans with Disabilities Act (i.e., QALYs were not derived from persons having disabilities; QALYs associated with disabled persons were underestimated in the Oregon approach). A more complete description of the Oregon example can be found in US Court of Appeals for the District of Columbia Circuit (1999). There, the Court did not judge the appropriateness of the DHHS denial of the Oregon approach, or the approach itself. However, they did cite the QALY as applied by Oregon as evidence of the existence of approaches that might be used by EPA as a starting point for its own use in deriving a method to determine a decision point for use in National Ambient Air Quality Standards decision making.

There are distinctions between the State of Oregon proposal, which affected the availability of medical treatment to identifiable individuals, and the policy uses contemplated by NCEA’s CRFM, which deals with public health impacts to individuals whose identity is not known (or perhaps knowable). Although the DWC doubts that this kind of legal-ethical objection would apply to decisions about investments in drinking water quality, it recommends that NCEA consider this issue and determine whether there are likely to be unanticipated legal-ethical dilemmas arising from the use of quality of life weighted longevity measures of disease impact.

f) The document should provide more information about both the variability and uncertainty in the QALY weights. It would be good to know, for example, (i) the variability of the QALY weights for the same conditions across individual respondents in the studies under consideration, (ii) the stability of the QALY weights for the same conditions from one study to another, (iii) the sensitivity of the QALY estimates to the set of conditions being estimated, (iv) the sensitivity to the method chosen for elicitation of the QALY weights (time tradeoff, standard gamble, subjective rating scales etc.), and v) whether there are genetic traits or environmental conditions within the human population that are potentially underestimated in the population samples that were studied.

A related issue is whether QALY weights for most anticipated EPA applications of the CFRM can be taken from the existing technical literature or whether they will need to be developed from field studies in the communities, or the specific populations, affected by the proposed interventions. Subsequent versions of the report should address this issue, which may substantially affect the feasibility and cost of implementing the CFRM. It would also be interesting to know whether EPA anticipates cases in which it would be appropriate to develop common metrics through, for example, focus group work with stakeholders.

g) There is a need for a better method for dealing with acute nonfatal disease. The assignment of QALY costs to health outcomes required a great many assumptions, combining results from a number of different studies and adapting results derived in one situation to a quite different situation. For example, the study used the QALY cost of a year in severe pain derived in the Hamilton, Ontario study to estimate the QALY cost of a bout of severe acute illness (such as diarrhea or vomiting), by dividing the one-year cost by 365 to get a one-day figure, and then multiplying by 14 for a two-week illness.

This procedure requires the assumption that pain produces "disutility" for a person at a constant rate. There is no empirical or theoretical basis for assuming proportionality, or for that matter, for assuming any alternative form such as marginal increase or decrease. In addition, the disparity in scale between a day and a year is very great; it is doubtful that the value to an individual of anything of a day's duration can be inferred from information about the value of experiencing that same thing for a full year. Some things that can be borne for a day (with little impact) can lead to despair if they must be endured for much longer periods. Other things, perhaps, we can get used to or can find ways to cushion ourselves from -- in which case the cost of a single day might be much greater than the average daily cost of a year.

In addition, the DWC believes that the method used to estimate QALYs for acute nonfatal illness is somewhat *ad hoc*. If such diseases account for a large share of the health impact -- as they do for many of the scenarios of the current case study -- the QALY assignments for this class of health effects may require further attention. There are at least two alternative approaches to this question. The first approach would involve a search of the literature for studies that address directly the tradeoff between acute nonfatal disease and more serious, long-term diseases, for which more reliable

QALY estimates already exist. If no such studies exist, It would be useful to do a comparative risk survey that directly pits against each other the two risks examined in the case study (acute GI disease and cancer after some period of latency). The survey results could be used to estimate directly the QALY cost of the various outcomes of these diseases, which could then be compared to the QALY costs estimated in the current study. A close concordance between the survey-based QALY costs and those estimated in the case study would help build confidence in the method used to calculate QALYs. If there were no close correspondence, then survey results would nonetheless help analysts develop better estimates of the QALY cost of acute nonfatal disease. A second option, would involve expression of acute nonfatal disease costs in monetary terms, possibly with the assistance of researchers in the Office of Policy.

h) Accurate portrayal of the severity of microbial and chemically induced disease states is a critical issues in the CRFM. For microbial agents, our knowledge of infection, morbidity and mortality provides a reasonably dependable point of departure for such determinations because the information can be developed directly from human experience. If the effects of chemicals are expressed in the probability of contracting a specific disease it is possible to parameterize similar issues in the context of QALYs. However, the information developed from animal toxicological studies is neither as complete nor as convertible to “states” of health that are as easily recognized by the average lay person. The influence of this on meaningful QALY development is unclear. This problem may be best illustrated by the following example. Assume that chemically-induced changes in circulating estradiol concentrations in an experimental animal lead to a variety of “adverse” outcomes of different severity. These effects would include a variety of symptoms easily identified in a humans that might be described as varying degrees of discomfort. Clearly, these effects should impact the QALY at some level, but they may not be apparent from animal data. Conversely sustained exposures to effective levels can lead to a variety of serious health effects related to impaired reproductive capacity, developmental toxicities, or cancer in a variety of organs. In these cases the QALY may be estimated with some accuracy, but may or may not be more severe than sustained symptomatology. The subtlety of depending upon indirect evidence from experimental animals to predict QALYs that depend heavily on human perception is a conceptual problem that was not directly addressed.

3.2.5 Step 4: Cost Evaluation

The final step is to evaluate the financial costs of the treatment technologies and the medical costs and estimate a cost effectiveness ratio. Is this an appropriate final step in the methodology? Are there other interim steps that should be evaluated?

The final step, evaluating the financial costs of the treatment technologies and estimating a cost-effectiveness ratio, is appropriate. Estimating a cost-effectiveness ratio is useful in evaluating financial costs of the treatment technologies relative to QALYs. It is important to recognize that the outcome of this analysis can be biased in either direction by including or excluding certain costs in the assessment. The advantage of the proposed approach is that, if properly done, it will make the costs included in any assessment quite explicit. This should be beneficial to those interested in Agency decision making.

Several obvious costs seem to have been omitted from the Agency's case study (apparently to simplify the presentation). Some of these will have to be included as the applications of the methodology are expanded. Many of these are discussed more fully in the portions of the charge related to the case study. However, the DWC thinks that several need to be discussed in the method itself:

- a) A decision was made to include only QALYs among the losses occasioned by disease, omitting the cost of health care and lost output. This was done to help keep the size of the document manageable, however, the result of omitting these costs is to cause the costs of the two technologies to seem to be the same. Given the high cost of cancer treatment, however, that assumption is questionable.
- b) Although the point is subtle, the current case study does not consider the economic costs associated with health care and lost productivity that propagate through the economy from the costs noted in "a)" above. In our meeting, EPA personnel indicated that future analyses would take this into account. It is also important that issues of access to/availability of health care be considered as part of these costs.
- c) The document does not consider the regulatory oversight costs of the Agency itself that are associated with alternative policy choices. There are likely different behavioral implications and therefore different costs associated with alternative regulatory prescriptions. In some cases these could be important additions to the cost evaluation.

4. THE CASE STUDY

This section of the report contains the Committee's general comments on the case study, including remarks on its limitations and on the appropriate handling of uncertainties associated with the method and the study. As noted in Section 2 of this report, the Agency directed a number of questions to the Committee on the case study. Although the Committee considers these questions to be of consequence and important to its review, it has placed them later in the document in Appendix A. There, detailed comments are directed at specific case study charge questions. The Committee has chosen this approach in order to give emphasis to its comments on the method. Just as the case study was conducted to illustrate the CRFM method, the Committee's review of the case study was conducted to determine how the procedures contained in the CRFM are implemented. In many instances, the results from the Committee's consideration of the study are generalized in its comments on the methodology itself which are contained in Section 3 of this report. Many of the comments in Section 4, and in Appendix A focus on shortcomings of the case study, therefore, they may appear to carry a more negative connotation than intended. Addressing the issues contained in the Committee's critique will be critical to further developing the methodology and to determining how generally the approach can be applied to the microbial and disinfectant byproduct area.

4.1 General Issues for the Case Study

4.1.1 Is the Case Study an Appropriate Application of the Framework Methodology?

Is this case study an appropriate application of this methodology? Are these initial

assumptions plausible for a hypothetical site? Are there other factors that should be considered within the scope of the case study?

The case study provided insights into how the Comparative Risk Framework Methodology (CRFM) would be applied, thus helping to demonstrate the potential of the methodology in formulating, structuring, and describing complex environmental problems for evaluation. The structure provided is more valuable than the calculated estimates generated because that framework makes explicit the factors that have been considered in assessing the risks and benefits associated with alternative drinking water treatment approaches. Thus, the study suggests that the CRFM could provide a good structural mechanism for improving discussions of complex environmental problems within the community of persons responsible for and interested in regulatory decision making.

As indicated in the charge, the case study was built around a set of circumstances that is likely to represent a particular local water system rather than a general set of circumstances. Consequently, the application to a very specific circumstance was more easily visualized than applications at the national level. However, it is unclear whether the intent was simply to model a single site—if so the assumptions generally seemed appropriate; to model sites more broadly—if so, this case may not be typical of water supplies in the nation as a whole; or to model a community where the issues entertained would be at the forefront of a regulatory decision—if so, the bromide contaminant may not be sufficiently represented in the scenarios presented to the Committee. The point is that a major variable in this type of analysis will be variability in the source water. The case study would have benefitted from some consideration of variations in source, because the effectiveness of competing technological fixes can be considerably different with variations in the water quality.

The case study ignores a number of important behavioral aspects associated with risk that should be part of the methodology. Several examples come to mind. The possibility of home interventions (e.g., the provision of point-of-use filters) raises behavioral issues that don't come up when dealing with the more customary centralized water treatment technologies. Some issues were mentioned at the meeting, including the possibility that the use of bottled water among AIDS sufferers is already very high possibly lessening the incremental gains estimated for home filtration; the incentive to return to tap water for those now using bottled water if filters are to be provided gratis; and the possible strong local pressure to expand use of home filtration for other sensitive populations (having sensitivities across a broad range) if the filters are regarded as a success.

EPA representatives at the February, 1999 DWC meeting indicated that the next step would be to develop an applications document that will include a collection of case studies. The Committee endorses the conduct of additional cases studies and suggests that certain studies might be especially useful. For example, an evaluation of the cost-effectiveness of conventional water treatment with filtration versus no filtration, or conventional water treatment with filtration vs. ozonation with no filtration, for large municipalities would be very useful. As one expands the scale from the local (as in this case study) to the national scale, the cost-effectiveness of adding filtration to those systems that now only disinfect would also be important to add to the evaluation.

4.1.2 Statistical Approaches and Uncertainty

Are the statistical methods and assumptions used to develop parameter distributions (Chapter 5) reasonable and correctly applied? Are the simulation procedures and sensitivity analysis appropriate for these data (Chapter 6)? Are there other statistical techniques that could be applied to improve the CRFM?

An important aspect of the CRFM is the approach used to characterize uncertainty. Uncertainty analysis can focus on "parameter uncertainty" as the case study does, or it could be expanded to deal with "model uncertainty." While the SAB understands the tendency to focus on parameter uncertainty, because it is relatively simple to handle in an "objective" and defensible manner, it is possible that failure to deal with model uncertainty could prove to be a critical flaw. In many cases, model uncertainty will dominate parameter uncertainty and efforts which ignore it could lead to a false sense of confidence in the results of the cost-effectiveness analysis, and serious distortions of research priorities. For these reasons, the SAB recommends that model uncertainties be clearly identified and addressed. As a minimum, the robustness of conclusions to different plausible "models" could be explored in sensitivity analyses. Also, formal elicitation of expert judgment could be used to begin to understand the extent of legitimate scientific debate and disagreement on key issues and to the impact of this on the cost-effectiveness of alternative control strategies.

In general, the members of the Committee felt that the assessment of parameter uncertainties for the comparative risk framework methodology appeared to be, for the most part, on the right track. However, the DWC identified some aspects of uncertainty and sensitivity analyses that were not identified or sufficiently discussed in the case study. The aspects that should be considered in the further development of this particular application are provided below:

a) Uncertainty Regarding the Methodology

- i) The case study should further explain the process used to propagate the uncertainty and variability of model parameters to approximate the quantitative uncertainty of the model output [the cost-effectiveness (CE) ratio in the case study]. The draft report does not discuss the methodology in sufficient detail to permit the reader to fully understand what was done and why it was done.
- ii) The Committee recommends that the report be revised to include at least a qualitative discussion of the possible magnitude of model uncertainty, and if resources permit, a quantitative assessment of model uncertainty should be described and implemented in the case study.
- iii) When conducting uncertainty analyses it is important to decide up front whether the assessment endpoint (the CE ratio in the case study) is considered to be some measure of the average for a specified population or a distribution over a population. If a distribution is being estimated, then both variability (over the population) and uncertainty of parameters are present. In that case the uncertainty analysis must be conducted in ways such that the uncertainty and variability components are separated and properly propagated. It appears that this separation of uncertainty and variability was not done in the case study. The DWC believes this problem should be addressed in future revisions of the case study and in other case studies that may be conducted.

The methods used for this purpose should be fully explained and consistent with the full description of the propagation methodology laid out in the general methodology section of the report.

iv) The credibility of uncertainty/sensitivity analysis results depends in large part on the credibility of the distributions used (selected) for the parameters. The rationale for some of the selected distributions was not provided in the case study. This omission should be corrected. Also, when it is difficult to see which distribution is most appropriate, the sensitivity of the predictive model results should be obtained by using alternative input distributions.

v) The *Cryptosporidium* data used in the case study provided an interesting example, but the analysis of the data was complicated by the fact that the data sets were censored, i.e., many samples were reported as being less than the detection limit (see Table 5-15 on page 5-35 in the EPA document). It appears from this table that the mean of the total oocysts present was computed using only the samples for which the counts were above the detection limit. This approach will tend to give a computed mean that is too large since the non-detects, presumably, have lower counts than the detects. Future revisions of the case study should consider using statistical methods that have been developed for estimating the mean of censored data sets. A paper that discusses some of these methods is: Helsel (1990).

vi) The uncertainty analysis in the case study did not consider the possibility of correlations among model parameters. The DWC recommends that any future revisions of the case study should incorporate methods for propagating the uncertainty of both uncorrelated and correlated parameters.

vii) The Committee believes that the case study should describe why the propagated uncertainty information about the model output (CE ratio) is useful and how it should be used by decision makers.

viii) Apparently, the case study used simple random sampling to generate multiple values for the uncertainty and variable parameters in the model. However, other methods of generating multiple values, such as Latin Hypercube sampling, can also be used. Latin Hypercube sampling is known to generate data sets that are more representative of the underlying distribution being sampled than can be achieved by simple random sampling with the same number of realizations. The authors of the case study should consider if Latin Hypercube sampling should be used instead of simple random sampling. The rationale for the selected method should be provided.

b) Uncertainties in the Public Health Outcomes of Alternative Treatments

i) The CRFM attempts to compare the different public health outcomes of particular treatments. In addition to uncertainties inherent with predicting the frequency of adverse health outcomes (e.g. dose-response relationships, extrapolations to low exposures, and secondary spread of infection), the present case study is complicated by a large number of factors that affect exposure. Central to this are uncertainties in the outcomes of treatment as they are compounded by uncertainties in the

characterization of the source water, treatment efficacy, robustness of the treatment train to process upset, and management of the treatment processes. The methods used to propagate these types of uncertainties into the final analysis should be clearly described in the case study.

ii) Subsequent case studies should consider the role of secondary transmission of waterborne infectious agents because for most microbial pathogens it is responsible for significant morbidity. Data can be obtained from outbreak investigations and CDC investigators.

iii) The uncertainty associated with extrapolating high-dose animal toxicity data to low-dose chronic exposures in humans is really not dealt with in the document. Doing so might be expecting too much of the current case study. The case study combines “loose” indicators of disease outcome that have uncertain implications for the probability of disease with individual byproducts having some available test results (problem is generally associated with effects on reproductive and developmental toxicities). A tendency to combine associations into assumptions that are applied to a group of chemicals injects a series of conservative assumptions one on top of another. That could tend to magnify the chemical risks to a point that the outcomes are unreasonable. In the present case study extrapolation was not a problem because the microbial risks were so overwhelming as to make the chemical risks relatively unimportant. However, circumstances can be envisioned where inappropriate extrapolation could lead to an inappropriate public health decision.

4.1.3 Case Study Limitations

Are the limitations of the case study appropriate and clearly presented? Are there other significant limitations not identified?

Many limitations in the case study were clearly presented in the Agency’s document. The case study did raise a lot of issues about the general approach that need future consideration. In addition, the limitations of some of the parameters selected were not noted in the document.

An important limitation is the use of “active” *Cryptosporidium* oocyst concentrations in water. This use may lead to an underestimation of risk. If drinking water is being evaluated, then the presence of any oocysts in the treated waters could mean that oocysts have in fact broken through the treatment system’s barriers. It may be more realistic to use the total number of oocysts when evaluating such risk.

In some respects the case study would have been more interesting if it had focused on contaminant levels where real risk management decisions might have to be made. For example, within the case study a comparison was made that indicated that central treatment might have been as cost-effective for protecting the vulnerable population as employment of point of use devices. While the analysis found a difference in cost-effectiveness that favored point-of-use devices, this difference was well within the boundaries of uncertainty. There were a variety of very unrealistic assumptions about the reliability of point-of-use devices and behaviors that affect their reliability. If these other factors had been taken into account, there might be a strong reason for favoring central treatment. Since central treatment would also

provide ancillary benefits in improved drinking water quality, the risk manager might find this an easier scenario to advance.

As the methodology evolves, attention should be directed to the question of scale. This local case study focusing on a relatively simple situation required many assumptions. These assumptions had little effect on the outcome of the analysis, because of the large differences in risks that could be attributable to *Cryptosporidium* and disinfectant byproducts under the scenarios defined. Some concern must be expressed that some of these assumptions may dampen the significance of important variables as the analyses progress to more diverse scenarios, focus on multiple locations, or attempt to deal with problems aggregated at the national level. It seems reasonably safe to conclude that the CRFM would be useful at the local level if the expertise necessary to conduct the analysis and correctly interpret the results is broadly available (especially in light of the assumptions made). At a regional level, there may be sufficient commonality in major variables (primarily source water issues) that a relatively small number of variations in the analysis could still make the approach practical. However, at the national level the application becomes much more abstract and, consequently, it is difficult to determine whether the approach would be effective.

5. RESULTS

5.1 Are the results presented appropriate (Chapter 6) and the interpretations of the results (Chapter 7) appropriate? Are there additional discussion items that should be presented in either Chapters 6 or 7.

The results section was appropriately structured. The additional data that go into the final analysis is presented in a way that is more trackable than presented in Chapter 5.

The DWC has made numerous comments about some of the assumptions that underlie these final calculations throughout this report. Some of these assumptions can have major impacts on the results of the case study.

Chapter 7 provided a generally good overview of what the case study might say to a local decision maker or a risk manager responsible for making decisions on MCLs. The DWC generally agrees with the potential utility of the CRFM at the local level. The DWC found the discussion on how this might be used by the national risk manager somewhat naive. All the case study really addressed was which of two interventions might be most cost-effective in preventing harm from waterborne *Cryptosporidium* to the general population and to a sensitive population. This analysis is very limited in that it depends heavily on local conditions (i.e. are there viable *Cryptosporidium* oocysts in the source water). DBP risks were unimportant in the analysis of the case because they were overwhelmed by the microbial risk. As a consequence, it is difficult to determine how this case study should influence the development of MCLs for DBPs or for microorganism. The first consideration that must be addressed in developing an MCL for an individual agent is whether the agent presents an unacceptable risk. Such determinations are more or less independent of other risks in the water. A second level consideration is how one might implement the achievement of that goal. As indicated in the introduction, the Office of Groundwater and Drinking Water must implement their program in the context of the Safe Drinking Water Act and the MCLGs and MCLs are the mechanism provided.

The statement made on page 7-14 implies that microbial and DBP threats to health from drinking water are competing risks. A more accurate statement is that microbial risks are large if water is from a vulnerable source. The first obligation of a national risk manager is that he/she cannot promulgate regulations that would compromise the delivery of water that is free of waterborne infectious agents. That protection must be stacked up against accurate assessments of probable risk from agents that are not as effectively dealt with by conventional drinking water treatment or source protection.

The next questions is: "Given that there is a suite of treatment processes that can eliminate or at least minimize waterborne infectious disease, which of these methods produces the lowest risk from DBPs at equivalent levels of utility?" At this stage it is essential to understand the local source water conditions in even more detail. However, it is improbable that a local water utility will be in a position to evaluate these risks for all possible DBPs. Therefore, there needs to be a set of standards that state unequivocally the levels above which no individual DBP should be allowed to occur. The real question that arises in this case is the burden of proof that such occurrence is actually harmful at the projected levels.

The utility of the CRFM at the national level may be in its ability to order information. One application is to identify research needs as suggested in section 7.4 of NCEA's document. An equally important application would be to help the national risk manager view the types of behaviors that an MCL might trigger on the part of the regulated community. For example, will the need to meet this MCL force a utility in the direction of an untested or poorly evaluated treatment alternative. The practical outcome of the first regulation of trihalomethanes in the U.S. was to increase the use of chloramines for disinfection. In retrospect, this probably caused little harm, but it caused the Agency to recognize that regulations could influence existing disinfection. That recognition was a major stimulus for research and is the technical basis behind much of the research that has been identified on the microbial and disinfectant byproduct problem in the last decade. Consequently, the DWC encourages NCEA to press forward with development of the CRFM as a valuable tool for displaying all the relevant factors that need to be considered in the development of regulations. Perhaps the most important refinement that could be made with respect to the microbial and disinfectant byproduct world would be to incorporate consideration of how socioeconomic factors will influence the response of the regulated community to a new regulation.

A specific problem in this section is the calculation of cost effectiveness ratios for ozone pretreatment for the AIDS subpopulation. The calculation of "average cost-effectiveness" of ozone pretreatment in the AIDS subpopulation (p. 6-41) is given as \$2.27 per QALY. This calculation is both irrelevant and misleading and should be removed. Cost-effectiveness is an incremental concept; it compares the effects and cost with and without an intervention. Sometimes the increments are very large indeed – an entire water treatment plant, for example. The comparisons here are not appropriate because there is no way to supply pretreatment to the AIDS subpopulation at \$23.37 per persons without also supplying it to everyone else in the service area at the same unit cost.

6. RESEARCH NEEDS

6.1 Are the research needs clearly defined? Are the research needs that are highlighted appropriate given the information in the Case Study? Are there other research needs that should be identified?

The sweeping nature of some of the assumptions made in the case study highlight the need to conduct research that can provide both qualitative and quantitative information for the assessments that are essential components of the CRFM. It is clear that the analysis of both the microbial and disinfectant byproduct risks were controlled by the assumptions made rather than by any quantitative data. As a consequence a valuable outgrowth of applying the CRFM to this problem is the critical research needs that it helped to identify. It also helps to make clear the great responsibility that accrues to analysts to ensure that the host of simplifying assumptions does not obscure the uncertainties in the analyses. With this in mind, a number of research needs are apparent. One category of need is specific to the continued development of the CRFM and another focuses on research specific to the problem being evaluated (for this case, research on microbial and disinfection byproducts).

It is essential that these research needs are seen as distinct from the basic science research in environmental microbiology, toxicology and epidemiology of waterborne diseases from disinfection byproducts and infectious agents. While it is certainly true that advances in the basic sciences may reduce uncertainty in estimation of health risks and in the identification of cost-effective treatment technologies, without a solid basis for comparative risk assessment much of this basic research will not achieve its true potential for affecting policy determinations and improving public health.

6.2 Research Needs to Support Continued Development of Comparative Risk Framework Methodology

- a) Much of the background for the application of QALYs depends upon issues involving trade-offs in medical practice or preventive medicine that potentially affect the same individual. Research should be performed that confirms valuations that the general public would assign to QALYs related to infectious disease morbidity and mortality relevant to environmental risk management.
- b) There is a need for a better method to handle acute nonfatal disease. The assignment of QALY costs to health outcomes in Chapters 5 and 6 required a great many assumptions, combining results from a number of different studies and adapting results derived in one situation to a quite different situation. For example, the study used the QALY cost of a year in severe pain derived in the Hamilton, Ontario study to estimate the QALY cost of a bout of severe acute illness (such as diarrhea or vomiting), by dividing the one-year cost by 365 to get a one-day figure, and then multiplying by 14 for a two-week illness. This procedure requires the assumption that pain produces "disutility" for a person at a constant rate. There is no empirical or theoretical basis for assuming proportionality, or for that matter, for assuming any alternative form such as marginal increase or decrease. This assumption is a matter for empirical investigation.
- c) It appeared to some members of the DWC that the values assigned to disease morbidity and mortality might depend on the individuals place in life and the extent to

which they have dealt with serious medical problems in the past. Some attempt should be made to validate the QALYs that are assigned for microbial and chemically-induced disease outcomes from several disparate points of view that could be represented in the U.S. population and for the population represented by the case study.

d) Methods for addressing model uncertainty in policy analysis.

6.3 Research Specific to the Problem Evaluated

a) Research is needed to better characterize risks that might arise from low level exposure to certain organisms and disinfectant byproducts. These research activities should first focus on those contaminants most likely to induce effects seen in epidemiological studies or that are particularly important in the decision process (e.g. bromate, viability of *Cryptosporidium* oocysts). This research should include efforts to clarify the relationships between disease outcomes in animals and humans.

b) The hazards related to *Cryptosporidium* exposure are well identified, and the authors make it a point to emphasize the idea that more data are needed within the context of the present case study. To deal with the total problem, data are needed not only for *Cryptosporidium* but for other microorganisms as well.

c) The possible survival and colonization of the distribution system by bacterial pathogens need to be studied if pathogens other than *Cryptosporidium* (which neither replicates in nor colonizes the distribution system) are to be included in future case studies. That possibility is one point in the model that was not taken into consideration, since only *Cryptosporidium* was being evaluated and *Cryptosporidium* will not grow or colonize.

d) The estimated cancer risks associated with disinfectant byproducts are not small compared to risks that have precipitated other regulatory actions that have been taken by the Agency. Consequently, as long as one can be certain that an option does not sacrifice microbial risks, the actual tradeoffs with different disinfectant strategies may continue to depend on comparative risks associated with disinfectant byproducts produced by each method into the foreseeable future. There must be a way to more efficiently guide research in this area in the future and the CRFM can help in this regard. The DWC suggests that this application of the CRFM should be pursued aggressively in the near term.

e) The research needs identified in Appendix A.4.9 may appear to be the general types of needs that are always expressed and are difficult to challenge. However, the needs are not specific and draw on a limited segment of the literature that is available on the toxicological effects of disinfectant byproducts.

i) The needs articulated continue to pursue questions that attempt to aggregate risks from disinfectant byproducts by class. In the case of the trihalomethanes, this issue has been answered in sufficient detail to know that this aggregation is not an appropriate approach. Toxics mechanisms of brominated trihalomethanes are distinct from those of chloroform. These disparate

mechanisms suggest different approaches for risk assessment based under the auspices of the proposed new guidelines. Similar dichotomization of mechanisms are surfacing with the haloacetic acids. It is time to conduct research on the properties which are conferred to disinfectant byproducts by bromine substitution.

ii) The research needs should reflect health concerns for potential sensitive populations in much more specific terms, both for microbes and DBPs. For example, several potential problems can be identified with the dihaloacetates. Given to metabolically competent animals, they are very rapidly metabolized and therefore relatively weak toxicants when dose is expressed as external dose. However, they are very potent systemically. If a significant fraction of the population does not express the glutathione-S-transferase zeta that is responsible for a very large fraction of dichloroacetate metabolism, segments of the population could be very sensitive to these compounds. Moreover, it is probable that these concerns extend to other halogenated organic acids produced in the chlorination of drinking water.

f) The Agency has been conducting research on the effects of exposure to mixtures of toxic chemicals for the last several years. However, it is also important to conduct research on the effects of exposure to mixtures of microorganisms. In contaminated water, it is likely that more than one type of pathogen will be present. It is not known whether simultaneous exposure to different pathogens would increase susceptibility to infection or whether the effects would simply be the sum of the risks for the individual organisms.

7. GENERAL

7.1 Do you have other specific recommendations for the improvement of this case study?

The DWC recommendations for improving the case study are found within the responses to questions in Appendix A. It is not necessary to recount those suggestions here; however, a few general recommendations are emphasized below:

- a) NCEA should develop more succinct guidelines for translating adverse health effects from animal experiments to human conditions. This guidance should also identify how mechanism or mode of action is to be used in these translations. These issues will be daunting if they must be revisited with each case study.

- b) The criteria used to determine if there is an incremental exposure due to a disinfectant treatment need to be much better established. The DWC is skeptical of the exposure information on several grounds. What was the basis for determining something was a disinfectant byproduct? How well has the distribution of DBP concentrations over time and space within a water system been characterized? How is that information to be aggregated as it is scaled up to the national level? Clearly, NCEA must take an active role in advising the Office of Water on how the data in the Information Collection Rule might be better formatted for purposes of risk assessment.

The historical data utilized in this case study are of limited quality and utility.

c) The CRFM must begin to consider source water characteristics as a major variable in these analyses if the methodology is to be expanded beyond consideration of a single water supply. Source water characteristics are important even where a single water source is being considered, because that has major impacts on the efficacy of various treatment technologies.

d) Microbial risks need to be considered in significantly greater depth. *Cryptosporidium* is not a good prototype for other waterborne pathogens for many reasons, including its resistance to disinfection, its behavior in the distribution system and minimal risk for secondary spread. It also is important to recognize that different classes of microbes have differing sensitivities to different interventions. There were too many assumptions made in the current case studies that utilized behaviors of infectious agents from completely different classes.

e) Individual susceptibility must start to play a bigger role in evaluations of DBPs as well as with microorganisms. Potential and real genetic susceptibilities must be considered as well as loosely grouped and not altogether generalized susceptibilities based on age, for example. The Committee recognized that the lack of data has made it impossible to consider these variables in any depth. However, NIEHS and NCI are undertaking efforts to examine the issues of susceptibility with much greater specificity. EPA needs to be ready to make use of that information as it becomes available.

APPENDIX A

Specific Comments on the Case Study

The Drinking Water Committee reviewed and has prepared comments on the case study both in response to charge questions from the Agency and because of the importance of its focal topic. Many of the comments go beyond the charge as it related to the limited case study presented. These added comments made by the Committee identify problems and issues that should be addressed when the application of the methodology has been expanded to consider the issues of microbial and chemical risk more broadly. The Committee felt it would be most useful to keep the expanded comments that related to particular questions within the section.

1. Engineering/Water Treatment

a. Treatment Scenarios and Cost

Are the choices of water treatment scenarios for evaluation relative to the “real world” scenarios appropriate? Do the construction and operation cost estimates adequately reflect real-world facilities and point-of-use device costs?

It is likely that the treatment scenarios in the case study are too limited to adequately reflect “real world” situations. EPA’s analytical focus on the median performance of treatment processes, and on comparing the most probable risk of chronic illness from exposure to chemicals and endemic levels of disease may be inappropriate. The approach being evaluated may be simply “doing the wrong thing” because it omits consideration of the occurrence of outbreaks. Historically, the technical literature has shown that most waterborne disease outbreaks have been caused by process upset/failure. Endemic disease is a reasonable hypothesis that may be difficult to prove. The distinction is important because reducing endemic disease leads one to promote treatment practices that improve removal whereas reducing outbreaks leads one to promote treatment practices that improve reliability. While these two objectives are not mutually exclusive, pursuing one does not necessarily accomplish the other. Therefore, the case study, and the general method for evaluating treatment alternatives should include the probability of system failure (e.g., the probability of failure of chlorine versus ozone; the failure probability of central treatment versus home treatment). The treatment failure issue can be illustrated by a number of situations. For example, ozone is seemingly a more complicated system requiring air or oxygen clean up prior to going into the ozonator. It would likely have a higher probability of failure than a gaseous and/or liquid chlorine feeding system.

Also the effect of colonization of reverse osmosis (RO) units by heterotrophic bacteria and the risk of opportunistic pathogens for immunocompromised individuals should be considered. Payment (1991) reported that families with RO units with high levels of bacterial growth were at increased risk of gastrointestinal symptoms.

Traditionally, the most commonly accepted approach to achieving reliability is by using multiple treatment barriers. For example, deep, protected groundwater supplies, which have little risk of contamination, are often treated with chlorination alone, conversely, for several decades it has been accepted practice to use both filtration and disinfection for any water

supply exposed to significant sewage contamination. If greater removal of bacteria and viruses were the only goal, the application of greater amounts of chlorine might suffice, but the inclusion of two independent treatment barriers reduces risk in two ways: first it greatly increases the probability that at least one barrier will be operating at all times, and second it does the job of reducing the pathogens both by removal and inactivation, making the treatment plant's performance more robust.

The Committee was asked to comment on the cost estimates included in the case study. Some informal comments were offered at the two meetings during which the DWC discussed the CRFM with NCEA staff. However, the DWC does not feel that it is the appropriate group to provide further comments on Agency assumptions in this area. Determining the adequacy of these analyses would require detailed engineering and cost estimation that the Committee judged to be beyond the scope of a DWC review.

b. Treatment System Efficacy

The Agency asked the Committee to evaluate the following:

- i. The accuracy of water treatment system efficacy estimates, including the assumed distribution, and**
- ii. The appropriateness and reasonableness of water treatment system assumptions concerning DBP concentrations in finished drinking water.**

This aspect of the case study is difficult to answer directly. The Committee primarily used the case study to identify shortcomings of the general methodology. While it is easy to identify shortcomings in this case study, it is not as clear how these limitations might reflect on a broader application of the methodology. The following comments simply identify some limitations that the Committee thinks would be problematic if they were carried forward in a real world analysis. In general they reflect elements that were left out of the case study, and which should not be omitted if a real world scenario were evaluated using the methodology.

- i) The case study is limited to, "...the comparison of alternative water treatment technologies and not to comparing changes in technological applications (e.g. changes in the levels of chlorination)." This limitation is important because the degree to which a particular technology is applied can have profound effects on both its ability to reduce pathogens and its ability to produce chemical byproducts. Besides chlorine, there are other examples. At low doses ozone does not remove *Cryptosporidium*, but at high doses it is more likely to produce unacceptable levels of bromate.
- ii) No attempt was made to consider the effect of the distribution system on either microbial contaminants or disinfectant byproducts. Indeed some of the most important decisions in protecting public health in the next few years will have to focus on the distribution system as a variable. Do we use chloramines to maintain better residuals and lower coliform levels or do we use free chlorine to provide better protection to the intrusion of pathogens at the risk of higher levels of known chemical byproducts? In the future it is conceivable that primary pathogen reduction might be accomplished by either membrane filtration or granular media filtration and UV irradiation. If the water produced is still high in DBP formation potential, choices in the distribution system may have huge impacts on chemical risks from DBP levels in the distribution system.

iii) The case study assumed that no malfunctions occur that cause performance to deviate from system specifications – such malfunctions are one of the major risk-factors where outbreaks are concerned. Moreover this assumption is connected to the decision to use *Cryptosporidium* as the target organism. Existing systems, if they work perfectly, are expected to control the other pathogens and yet the discussion of outbreaks makes it clear that *Cryptosporidium* is not the only organism implicated.

iv) There are concerns that relate to the actual application of the CRFM. It could be a tool that is utilized for evaluating individual water supplies. On the other hand, it may be intended for broader national analyses of the impacts of various regulatory strategies. If the latter case is intended, the large and variable impact that source waters have is not taken into account in the case study. The Total Organic Carbon (TOC) in the water treated in the case study pilot plant is low, 1.76 mg/L during the ozonation study and 1.98 mg/L during the chlorination study. As a result, the DBPs formed in the pilot plant and used as inputs in the case study are low compared to many plants in the U.S.A. Bromide is also low, so that bromate production by ozone is not particularly high. As a result, the QALYs associated with DBPs are low in the study. Because the TOC is low, the applied ozone dose is low [the water has a low "ozone demand"]. This low demand results in lower costs for the ozonation process. The low TOC and low ozone dose also should result in a lower than typical production of biodegradable organic matter in the plant, simplifying the design and operation of the biofilter plant following the "pre-ozone" application. These may be "key parameters" for which a sensitivity analyses might provide some insight into implications they may have on a larger scale.

v) Baseline Technology - The conventional filtration plant is given credit for two logs of removal of *Cryptosporidium*, in accordance with the newly revised Surface Water Treatment Rule (SWTR) but less than suggested by the results observed in the pilot plant used for the case study (Table A.1-2, p. A-16). In contrast, ozonation is given credit for 0.5 to 1.5 logs of removal of *Cryptosporidium*, even though the pilot plant data only support 0.5 logs of removal [note also that the pilot plant operated at about 27 degrees C, a temperature that should favor inactivation]. These data indicate that the effectiveness of *Cryptosporidium* removal by the conventional plant was probably underestimated while the effectiveness of *Cryptosporidium* inactivation by ozonation was probably overestimated. This assumption could overemphasize the value of ozonation and lead to overestimates of the exposure of the general public and the AIDS population to *Cryptosporidium* when ozone is not used. Presumably, the QALYs would also change. These baseline assumptions may be "key parameters" and should have been subjected to a sensitivity analysis.

vi) Point of entry/use treatment devices are assumed to achieve 100% removal of *Cryptosporidium* all of the time. This assumption is not reasonable. Units will leak organisms depending on the quality of their materials, the assembly design and the quality of assembly and testing. Very small leaks around seals and fittings can result in significant contamination. More importantly, these units are not operated with supervision and indications are that they will often remain in place for some time after they fail. Thus, the failure rate is likely to be much higher than for central treatment.

vii) The case study pointed out that the home filtration technology is cost-effective even after implementation of the pre-ozonation technology. It is not clear why home filtration is considered to be the “marginal” technology in this analysis. The usual rule is to apply interventions in the order of their cost-effectiveness, which in this case would require the implementation of home filtration first followed by pre-ozonation if it could pass the cost-effectiveness test. That is, based on the costs and benefits as given, the two technologies are implemented in the wrong order.

2. Risk Characterization

a. Cancer Risks

Overall, are the cancer risk estimates based on exposures to mixtures of DBPs plausible and reasonable given the goals and limitations of the case study? Are the uncertainties presented in the assessments of DBP mixtures adequately described and characterized? Are there significant uncertainties that are not identified? Are the presentations and discussions of variability appropriate?

The report described risk characterization as the evaluation and integration of major scientific evidence and “bottom-line” results from hazard identification, dose-response assessment and exposure assessment. Toxicological data were utilized in a more or less conventional manner for those chemicals for which data actually existed. However, by far the greater number of compounds considered were evaluated were chemicals for which no toxicological data existed for important endpoints. These predictions for chemicals for which no data exist were based on predictions by a computer model known as TOPKAT. While TOPKAT is an established program, one can gain very little mechanistic insight as to why some chemicals were considered to be positive and others were considered negative for carcinogenic potential. Recent evaluations of TOPKAT and similar programs have shown that expert judgment appears to be more dependable than their modeled predictions once the chemicals are actually tested (Ashby and Tennant, 1994).

The document did not identify the evidence supporting inclusion of some case study chemicals as DBPs, except for noting that the Office of Water identified them as such. Some of these chemicals are not familiar to those on the panel who are, themselves, familiar with DBPs. Moreover, it is virtually certain that meaningful exposure data do not exist for a large majority of these compounds. As a consequence a significant portion of the toxicological risks projected are based upon chemicals with an uncertain status as DBPs, virtually no specific toxicological information, and little available exposure data. The document is unclear on how much these compounds influenced the overall toxicological risk estimates.

Further, the predicted carcinogenic risk from these analyses were summed using a response addition model and all were assumed to increase the incidence of bladder, colon and colorectal cancer in the human population. These were the target organs identified in epidemiological investigations; however, they were not the endpoints seen in the animal toxicology studies. The epidemiological studies were discounted as providing sufficient information to project risks; however, it was apparently used to distinguish cancer illness from cancer death estimates.

Essentially the same approach was used for estimating the aggregate risks for non-

cancer endpoints. Several aspects of this approach vary from traditional EPA evaluation approaches. The body-weight to surface area correction was used in cancer risk assessment but not the non-cancer assessment and the response addition model was used to sum the activities of chemicals for which thresholds could not be excluded. Since these are changes from the usual practice, some justification for these departures is appropriate.

The case study needs a consolidated section on risk characterization and that section should discuss the limitations and uncertainties. In general, mention of the uncertainties in the method were scattered throughout the text of the document and some of the more subtle departures from usual procedures described above, were not discussed at all. There was a very general discussion of uncertainty on page xxvii of the executive summary. However, it was very difficult to appreciate specific uncertainties across the different sections of the case study itself. The DWC strongly suggests that these are critical issues for the risk manager and must be brought together in a more easily identifiable way, perhaps in a separate section. In that section, a table listing the uncertainties should distinguish between uncertainties due to the lack of data, uncertainties in prediction (e.g. structure activity analyses), and uncertainties that relate to quantification of risks. The latter issue is usually grounded in science, but virtually always steps into the policy arena. Some sense of this transition is provided in Appendix A, but it is not explicit. Further, a discussion of how these limitations and uncertainties might affect the estimates of risk should be included in the text associated with the list.

The approach taken in the case study did illustrate some tools that could be applied in a comparative risk analysis; however, the DWC has significant concerns about the validity of cancer risk estimates that arose from the analysis. The reasons for this concern derive largely from the uncertainty of the data base and the assumptions that were needed to use it in the case study.

In each analysis, there will be a need to derive a best estimate with “confidence intervals” that capture the uncertainties in the estimate (both qualitative and quantitative). Discussions of variability in the DBP cancer risk estimates were not included as a separate issue. There were general discussions of the possibility that children or the aged might be more susceptible. A more specific discussion of variability was included for the exposure variable. A major issue of variability that was not discussed adequately is the variation in susceptibility in human populations. There are sufficient data for some of the DBPs to project some determinants of that susceptibility. Of course, there is no basis for discussing variable responses for the vast majority of chemicals considered.

b. Microbiological Risks

i) Overall, are the risk estimates based on exposures to the pathogen plausible and reasonable given the goals and limitations of the case study? Risks from only a single type of microorganism were assessed in the case study; was this an appropriate selection. If so, are there other organisms which should be evaluated with this approach in future applications of the CRFM and why?

Are the uncertainties presented in the assessments of pathogens

adequately described and characterized? Are there significant uncertainties that are not identified? Are the presentations and discussions of variability appropriate?

The selection of *Cryptosporidium* for this first case study of the CRFM was an obvious choice. This organism is an important pathogen that is well studied, has been shown to be transmitted by water, causes disease in both special populations and the general public, does not replicate in the environment, and is very resistant to removal and inactivation by current treatment techniques.

The case study assumes that an individual can be reinfected with *Cryptosporidium* every twelve weeks. This assumption corresponds to no immunity. Further, during the interaction with the Agency, NCEA representatives noted that this assumption was based on studies of Schistosomiasis. This should be explained more fully in the document. In fact, studies by Chappell (1998) indicate that there may be some protective effect as a result of exposure to *Cryptosporidium*. This study is further discussed in subsection b.ii. below.

Eventually, the CRFM will have to include consideration of other pathogens where the variables will be more numerous and complex. A major issue will be the fate of pathogens in the distribution system. Most pathogens can be inactivated during treatment, but some can remain infective and reach the distribution system. Some bacterial pathogens can even colonize the distribution system. This was not considered in the case study, and it is not an issue as long as *Cryptosporidium* is the only pathogen considered because it will not colonize the distribution system. However, this factor should be considered in future case studies, especially if bacterial pathogens are considered in the analysis. As application of the method is expanded, it will be necessary to consider that there is not a consistent ranking in the susceptibilities of waterborne pathogens (e.g., viruses, bacteria, other protozoa) to disinfection strategies (e.g., chlorine, ozone, UV light). This will also have to be taken into consideration in future case studies.

Weighing the benefits and costs of two interventions when multiple waterborne pathogen are being considered will require a more sophisticated data base. If there is evidence that a specific intervention, such as ozonation, will reduce the risks from several waterborne pathogens, then all such reductions should be included in the overall consideration of the intervention. Not to do so leads to an underestimate of the effect of the intervention.

ii) Are the identified hazards related to exposure to the protozoan parasite *Cryptosporidium parvum* appropriate?

-Are the probabilities of infection and disease conditional on exposure appropriate for the two populations considered (i.e., the general population and the AIDS subpopulation)? Have all of the possible outcome categories been identified?

-Is the scientific basis of the dose-response model for predicting pathogen risks well-founded and adequately described and supported?

-Should other subpopulations (e.g., the elderly) be evaluated in future applications of the CRFM?

Hazards from *Cryptosporidium* exposure are discussed and the authors emphasize the need for more data in this regard. However, the case study and the methodology focus entirely on endemic hazards from *Cryptosporidium* (by implication other microbial contaminants would be handled in the same manner). There is little recognition in the document that indicates that endemic hazard is less well established in drinking water than is the occurrence of sporadic outbreak of infections.

The distinction between sporadic outbreaks and endemic disease is important because the treatment strategies appropriate to these two circumstances differ significantly. This is also not recognized in the case study. While there may be some reason to be concerned with endemic disease, it is not appropriate to ignore established hazards that occur as outbreaks.

The dose-response model for the prediction of risks associated with *Cryptosporidium* is based upon studies in healthy volunteers. There are other populations that need to be considered in future case studies. It appears that other subpopulations, such as the elderly, may be at increased risk from *Cryptosporidium*. It has been suggested that the mortality ratio for individuals in nursing homes may be as high as 50%. If, indeed, this mortality risk is present, it would be a significant consideration in the risk assessment calculations.

Infectivity of *Cryptosporidium* is another important factor for this risk (as is shown in the sensitivity analysis). Significant sources of uncertainty exist for these estimates and the following issues should be considered:

aa) Estimates of the infectivity of *Cryptosporidium* should be based directly on analysis of the human challenge studies conducted by Chappell (1998). Do not rely on the interpretation by Perz et al (1998). There appears to be a large range of dose-response associated with various strains of *Cryptosporidium*. Dr. Chappell's dose-response research indicates that the ID50 ranged from around 9 oocysts for the TAMU strain to 132 for the Iowa strain to 1100 oocysts for the UCP strain. To date, they have only tested genotype 2 *Cryptosporidium* strains (animal strains) and not genotype 1 (human) strains.

bb) What is the effect of pre-existing anti-Crypto antibodies on host susceptibility to infection and what proportion of the population (stratified by age) have antibodies to *Cryptosporidium*? Dr. Chappell observed that antibody positive volunteers had an ID50 of 1880 oocysts (Iowa strain), which suggests that there may be some protective effect that is overcome at high doses. Antibody positive subjects had more severe illness and longer duration of illness, but fewer of these subjects shed oocysts at detectable levels, and those who did shed oocysts did so at lower concentrations. The implications of this for secondary transmission should be considered.

cc) What are the probabilities of reinfection? The case study assumes that an individual can be reinfected as frequently as every 12 weeks (page 5-47). The reference for this is Hurst et al. (1996) but these authors are not infectious disease clinicians or epidemiologists. The primary source of this information should be explored. EPA staff explained at the February 1999 meeting that data on schistosomiasis was used to estimate probability of reinfection of *Cryptosporidium*. It is not appropriate to extrapolate from a helminth infection with a very different life cycle

and pathogenesis to a protozoan infection. The data from Dr. Chappell's rechallenge experiments should be examined, and the opinion of Centers for Disease Control (CDC) experts in the Parasitic Diseases Branch should be solicited. Because there are very limited data on this issue, it should be treated as a source of uncertainty.

dd) In calculating risk, EPA assumed that the probability of infection depends on the number of organisms consumed in a 12-week period. In doing this, EPA assumed that the probability of infection is the same whether one is exposed to X organisms in a single dose or in X/84 organisms in 84 daily doses. There are no data to support this assumption. The resulting risk may be an over- or under-estimation of the actual risk.

ee) What are the probabilities of mild illness, moderate to severe illness and death given infection for immunocompetent and immunocompromised hosts? Again, it is important not to rely solely on the analyses of Perz et al. (1998). Investigators at CDC should be contacted regarding what they have observed in outbreak investigations and studies of *Cryptosporidium* infections in immunocompromised populations. CDC should also have information about the infectious dose for immunocompromised individuals compared to immunocompetent individuals. Infectious disease clinicians differ in their opinion on this matter. Some believe that, if stomach acidity is in a normal range, then the infectious dose would be the same for immunocompromised and immunocompetent individuals. Others clinicians believe that the infectious dose is lower for AIDS patients.

3. Chemical Dose-Response Assessment

The goal of both DBP and microbial risk estimation is to provide realistic, central tendency risk estimates that can be used to compare the treatment technologies in question. EPA chose to use a response addition model (as described in the 1986 US EPA Mixtures Guidelines); see also Hertzberg et al. (in press) based on the lack of data on the mechanism of action of the DBP. This is a mixtures approach to addressing DBP toxicity. Other approaches, for example, another mixtures approach such as dose addition or an assessment of individual chemical constituents of the mixture, could have been developed. It should be noted that, under the conditions set forth in the case study, DBP risk is significantly smaller than the risks posed by exposure to *Cryptosporidium*.

a. Is the discussion of mechanistic toxicity data for the individual chemicals and/or mixtures appropriate?

The Committee did not believe that reproductive and developmental toxicity data were handled correctly in the case study. Part of the Committee's concerns were expressed in section 3.2.3. In addition, the use of the response addition model, with its assumption of no threshold, is not consistent with Agency policy for the assessment of these endpoints. As noted, this makes little difference for the present case study. However, once this type of simplification has been used, it may set a precedent that is likely not to be acceptable where these hazards are more critical.

The current document provides relatively little description, analysis, or summary defining the mechanism of action for the individual DBPs or for DBP mixtures. The substantial expansion of discussion to include defining the mechanism of action would probably serve little useful purpose in this case study, as knowledge of mechanism of action did not appear to influence the model inputs or the interpretation of the model results (except for chloroform).

However, for a variety of reasons, the Committee felt that it would have been useful to group DBPs by presumed modes of action in a table (including an “unknown” category). First, as discussed in the next section, this information should underpin the model(s) used for estimating risks from mixtures. Second, the Agency is currently developing a document, *Guidance for Identifying Pesticides and Other Chemicals That Have a Common Mechanism of Toxicity*. To include this information would provide some consistency with other Agency efforts. Third, such a summary might provide a useful starting point for a more explicit rationale for using Quantitative Structure Activity Relationship (QSAR) predictions of the activity of compounds for which there were no toxicological data. Essentially, these chemicals all appeared to be treated as non-threshold toxicants, whereas chemicals with data frequently appeared to be treated as if there was a threshold [e.g. bromodichloromethane (BDCM) effects on developmental toxicology].

b. Dose-response analysis of individual DBPs (point added by the DWC)

The individual DBP dose-responses were modeled by a linearized multi-stage model with a threshold parameter using toxicity data from the Integrated Risk Information System (IRIS). The use of this model and the IRIS data is appropriate and justified, as a starting point. However, IRIS typically has not contained sufficient information to allow uncertainty to be evaluated and propagated by the model. The case study failed to provide an adequate assessment of the strengths and weaknesses of the studies used. For example, the route of DBP exposure used for various studies, corn oil gavage vs drinking water, does not appear to have entered into the evaluation of the data. In addition, several data sets failed to converge. The case study contained little evaluation or analysis of the adequacy of the linearized multi-stage model to adequately fit the IRIS data sets, nor was there an evaluation or analysis of the adequacy of the IRIS data for the dose-response model used. In this sense, the general methodology and the case study do not provide adequate direction to ensure consistency in application of the approach to new case studies. The DWC also is concerned that the IRIS data set is incomplete, may lack quality control, or be comprised of obsolete data. Thus the use of IRIS data should be supplanted or supplemented by primary data when appropriate.

c. Appropriateness of assumptions, techniques and models used in the estimation of risks posed by DBP mixtures, including both statistical and biological considerations.

The choice of a response additivity model to predict the health effects of the DBPs is an appropriate approach for some types of toxicological effects. The response addition approach is appropriate for summing risks for certain mechanisms of carcinogenesis, for example. Clearly NCEA understood this because it chose not to treat chloroform risk as response additive. To be consistent, the possibility of other sublinear dose-response relationships should have been entertained more explicitly. The selection of the response addition model for reproductive and developmental toxicities is inconsistent with previous

Agency policy in the drinking water area. For chemicals where thresholds are usually assumed the dose-addition model should be used. Applying the response addition model in this “exceptional way” should be specifically justified.

In the context of this case study, the selection of response addition vs. dose addition has little impact on the outcome. However, situations defined by future case studies or applications of the methodology may include chemical exposure scenarios for which experimental or epidemiological data support interactive toxicity (synergism or antagonism). Neither the general methods, nor the case study section, defined criteria or approaches for the inclusion or exclusion of data that support such interactive effects. The assumption of independent action as a requirement for using the response addition model is not one that deals effectively with the biology of interactive effects. There must be a methodology for integrating such effects into the process, even if they are relatively rare at low doses.

Other assumptions made in the case study are probably not appropriate for the analysis. For example, lumping all developmental toxicities into the highly dependent category. Clearly, the degree to which this lumping is valid depends on the nature of the birth defect. It is not established that minor effects, such as decreased crown-to-rump length, actually lead to some level of dependence. Such measures are useful screening assays, but they are not predictive for more serious effects, and make the transition to QALYs very difficult if not impossible. An assumption of dependency might be more reasonable for some other effects such as valvular defects in the heart that have been detected with many DBPs. More specific criteria should be developed to provide guidance for extrapolation of developmental toxicities than simply conversion from a continuous to a stochastic response.

The General Methodology section should also consider defining strategies to conceptually address the cumulative risk of exposure to multiple pathogens.

d. Usefulness of the mixtures risk assessment approach for comparing health risks across drinking water treatment scenarios and levels of DBPs.

The use of the mixtures risk assessment approach was valuable in that it provided an illustration of how the framework might be applied. This methodology has its roots in statistical rather than mechanistic considerations and as a result provides a somewhat distorted view of the possibilities that exist for interactions. Independence of action is an assumption for response addition, primarily because it is assumed that differing effects are not related and at low dose will be randomly expressed in the population. Synergy arises from independent action as well, but affect function through converging and perhaps redundant pathways. This requires distinct mechanisms of action, but sets up the situation for an amplified severity of effects and observation of effects where none were previously observed.

Consequently, it is important that the underlying mechanisms by which effects are produced be explicitly identified whenever possible. This identification becomes even more important as QSAR is used to predict toxicological responses. Inherently, QSAR predictions are based in large part on common mechanisms. A difficulty encountered in the report is that in many cases it was tacitly accepted that chemicals in the same classes acted by the same mechanism. There is certainly more than one mechanism involved in trihalomethane and haloacetic acid's effects. There has been at least one report of synergy between low doses of

dichloroacetate and trichloroacetate where the effects of high doses of both byproducts results in inhibition of one another's carcinogenic effects (Pereira and Phelps, 1996) clearly the two compounds produce liver cancer by distinct mechanisms (Stauber et al., 1998). It is extremely important that these differences in mechanism be kept track of in comparative risk scenarios where projections of effects are extended over a large number of chemicals for which descriptive toxicological data are not even available.

e. Appropriateness of assessment of risks posed by the unidentified halogenated fraction (TOX).

The DWC has reservations about extrapolating risks associated with those DBPs that have been extensively studied to the much greater number of DBPs that have not been studied. First, the compounds that have been studied do not even come from the same chemical classes as those that have not been studied. Second, within the classes that have been studied there are significantly different potencies that seem to be related to differences in the mechanisms of metabolic activation (e.g., brominated versus chlorinated trihalomethanes). Third there are diverse types of effects produced (e.g., cancer, reproductive toxicity, developmental effects) which may or may not have common mechanisms. Fourth, within the context of DBPs that have been identified as carcinogens, neither the target organs, nor the potencies predicted, approach those projected by analyses based on epidemiology studies that have been published by a number of investigators. Part of this difference might be explained by the relative lack of data on most DBPs, but there are also other potential explanations. The epidemiology data may be in error, there may be interactions that have not been appropriately identified, or humans may simply respond differently. These all must be treated as uncertainties. While the inclusion of this projection provides some conservativeness to the project, it does not diminish the level of uncertainty. The CRFM does not provide guidance on how to handle the uncertainties associated with these projections.

Having made the general point above, the Committee notes that extrapolation of risk from known DBPs to the unidentified halogenated fraction (TOX) was not a critical issue in the context of the present case study because the health impacts were dominated by the overwhelming contribution of *Cryptosporidium*. That is to say, disinfection byproducts, in general, had little impact on the method's results and conclusion. However, analysis of other cases will require site-specific estimates of the likely make-up of the byproducts that will be formed. The byproducts produced by ozone and chlorine dioxide are primarily non-halogenated or are not organic and are not captured within the TOX. Moreover, source water quality makes major differences in both the type and extent of byproduct formation. Thus, the extrapolation of the health risk from DBPs to the TOX in this case scenario must be recognized as an exercise that is specific to circumstances of this particular case and, therefore, has limited, if any, application to other case studies.

f. Have the appropriate DBPs been addressed? Cancer that results from chloroform was not addressed in the analysis because of the evidence (Drinking Water NODA, EPA, 1998) indicating that this chemical's mechanism of action exhibits a threshold. Is this an appropriate decision?

One must start with the disinfectant byproducts for which sufficient toxicological information is available to make an approximate estimate of cancer risk. The analysis

excluded chloroform because of conclusions that were published in the 1998 EPA Notice Of Data Availability (NODA). On the basis of that conclusion, it was appropriate to drop it from the analysis because the concentrations seen in the water consumed were below the MCLG put forward in that NODA. However, since that time, the Agency has modified its opinion on how chloroform will be handled in its regulatory rule making and the SAB must reserve judgment on specific chloroform issues until it conducts the review discussed in the EPA final Stage 1 rulemaking for DBPs.

The removal of chloroform from consideration in the case study raises an issue of how confined the CRFM would be by Agency policy. As indicated in previous sections, there are data that suggest modes of action among various carcinogenic DBPs that might warrant similar treatment under the proposed new cancer risk assessment guidelines. Therefore, one has the dilemma of whether the analysis can go ahead on the basis of current science or whether it must await wider policy decisions to be made, a process that can be very slow. It would be inappropriate for the DWC to prejudge Agency policy in this kind of exercise. However, we can suggest that it would have been useful to illustrate the relative impact of linear and non-linear extrapolations to low dose in this case study. As it turned out, of course, the disinfectant byproducts contributed little to the risk trade-off in this illustrative case study and such an effort would have been only an academic exercise. In a case study aimed more specifically at the relative risks associated with byproducts from different disinfectants, these questions would have been critical.

g. For future applications of the CRFM there is a potential interest in using the results of epidemiologic studies, please provide suggestions for the use of the epidemiologic data in the future applications.

Epidemiological evidence that is consistent across multiple studies can provide a basis for estimating the health risks associated with waterborne disease. In these circumstances, the data should be used to the extent that they are dependable (e.g. data may provide strong evidence for hazard identification, but not be appropriate for quantitative estimates of risk) When epidemiological data are uncertain, it is important to rely on toxicological information. A strong case can be made if the toxicological and epidemiological data present a consistent picture from a qualitative and quantitative point of view. As noted, however, a level of consistency from the two data sources has yet to be achieved.

There are some shortcomings of epidemiology that need to be recognized. Some epidemiological data support the hypothesis that endemic risks from microbes in drinking water exist. However, the data do not indicate if the risks are governed by the source water or by the distribution system. Without more refined information, it is difficult to translate these data into a strategy for mitigation. Clearly, more quantitative approaches are needed to make these data useful within the context of the methodology.

In a similar vein, the epidemiological evidence associating disinfected drinking water with cancer and spontaneous abortions is interesting. The findings suggest that there are as yet unresolved issues about disinfectants and their byproducts. To accept these data as indicative of a problem raises the question of which treatment strategies would mitigate such risks. These risks were associated with chlorination; however, it cannot be stated with confidence that these risks would not be found with alternative disinfection strategies. Thus,

these data serve as an incentive to better characterize the complex byproduct mixtures of chemicals resulting from chlorination in subsequent epidemiological studies. Also, toxicological studies that are designed to identify the cause(s) of the increased risk are needed. The CRFM can be a powerful tool for identifying shortcomings of this type in the available data. That would provide the Agency with a much better appreciation of how it should spend its research dollars to resolve issues with more specificity than has been possible in the past.

It is also important to recognize that epidemiological investigations provide other useful data for use in the CRFM. Specifically, waterborne disease outbreak investigations provide useful information on risk factors associated with outbreaks, susceptible host populations, incubation period, clinical characteristics of infection, secondary transmission of infectious agents, effective interventions and the ability of disease surveillance systems to recognize outbreaks.

4. Exposure

a. General issues related to the evaluation of exposures

The accuracy of exposure assessment in this case study depends on information on the occurrence of the microorganisms or disinfection byproducts of interest and on estimates of individual exposure to tapwater via ingestion, inhalation and dermal contact. For both disinfection byproducts and microbial pathogens, surrogate parameters are often measured instead of the occurrence of the actual pathogenic organism or hazardous DBP. Assumptions must be made about the accuracy of the measured water quality parameter (sensitivity and specificity) and how well this parameter serves as an indicator for the presence of the actual hazardous waterborne substance.

In general, the assumption of steady state conditions in source water and treatment efficacy (e.g., average annual exposures) departs significantly from real conditions in many treatment plants (especially surface water sources) for both microbial contamination and disinfection byproducts. Source water variation and variation within the water distribution system can significantly affect concentrations and types of microbial pathogens and disinfection byproducts. Most microbial waterborne disease outbreaks have been associated with a spike of microorganisms in the distribution system because of spikes of contaminants entering the source water (spills, spring melts, heavy rainfall, etc.), failures at treatment plants, or failures in the distribution system. Potential failure of the reverse osmosis units and problems with monitoring their performance also need to be considered in the case study. It may be more realistic for the model to identify distributions of exposure that take all of these variables into account rather than average annual exposures.

Concentration data for individual DBPs in the case study were recalculated from sampling data assuming a log normal probability distribution function and substituting half the detection limit for non-detects instead of zero. The Committee discussed "The Use of Censored Data" in section 6.2 in its July 19, 1995 SAB "Review of Issues Related to the Regulation of Arsenic in Drinking Water." The comments made then are applicable to exposure calculations with censored DBP data as well.

Source water variations can result in significant changes from one treatment plant to another in identifying the potential concentration of *Cryptosporidium* oocysts in the treated water. The most recent meeting conducted by EPA with a panel of statistical experts on clearly very preliminary data from the Information Collection Rule showed at best a 20% detection of oocysts in the treated water of those sites analyzed, with as low as 5% of the sites showing positive oocysts in their treated water. Site variability is extremely critical as is being able to evaluate the risks of ozone and/or membrane filters at the individual home.

The case study used recent *Cryptosporidium* occurrence data from LeChevallier in Trenton, NJ. The document uses the term “active” oocyst, which is confusing. It is almost a contradiction in terms. The determination of whether oocysts are likely to be infective is dealt with in operational rather than absolute terms. The terms conventionally used to describe such oocysts are: total, viable (based on dye permeability?) and infective (based on cell culture or mouse assay). These definitions should be clarified in future case studies and discussions of the method itself.

The Committee suggests that an upper bound on risks might be estimated by using total number of oocysts instead of total number of “active” or “viable” oocysts. Given that the methods to detect *Cryptosporidium* in drinking water are insensitive, any detection of oocysts in treated water should be used as an indication of the possible presence of infectious oocysts in the water and a potential threat to human health. However, it must be recognized that the most probable estimate of risk depends upon the viability of oocysts and their infectivity to humans.

The effect of colonization of reverse osmosis units by heterotrophic bacteria and the risk of opportunistic pathogens for immunocompromised individuals should be considered. See the paper by Payment et al. (1991). Payment reported that families with RO units with high levels of bacterial growth were at increased risk of gastrointestinal symptoms. How do you monitor bacterial colonization and water quality from these units?

b. Is the distribution of drinking water consumption rates appropriate?

The DWC was unable to determine the exact distribution of drinking water consumption rates used, although there was a significant amount of discussion related to the topic in the case study. The report (Section 5.2) cited the US EPA Exposure Factors Handbook (1997) in providing tap water consumption rates for the general population. The source of the data was Ershow and Cantor (1991). The rates were listed according to age groups. Then the report provided age-weighted averages of the values to approximate consumption by 5-yr increments. The data from Canadian Minister of Health and Welfare (1981) were used to correct for the unheated tap water consumption rates. The Perz et al. (1998) study was used to derive the percentage (70%) of unheated tapwater for the AIDS population compared to the general population, while total consumption was assumed to be the same. The Ershow et al. (1991) data were used as a basis for having no adjustment made to reflect any potential differences between pregnant women and the general population.

At this point in the Agency report (end of Section 5.2), before moving on to the discussion of DBPs, or in Section 5.3.2, it would be useful to clearly specify the exact consumption values used for deriving the risk estimates in the later part of the report, along

with information on the associated parameters (e.g., age-weighted adjusted, percentile, total water consumption). In Section 5.3.1, tap water consumption is represented by Y in the equation which described the response addition model used in the case study. Section 5.3.2 described the assumptions made for tap water consumption. But it seems that there is no clear indication as to the final values used. For example, if the values are from Table 5-2 as noted in Section 5.3.5 then it should be indicated right after the discussion of the data sources for consumption rates. The correct values are needed for verification of the results.

A related question is whether the water consumption rates used in the case study are the same consumption rates used by U.S. EPA in its programs and are they consistent across EPA programs? The Office of Groundwater and Drinking Water is developing statistical treatments of data gathered from the US Department of Agriculture's Continuing Survey of Food Intake by Individuals that may be of use in future application of the CRFM to drinking water problems.

c. Is the identification of the fraction of unheated drinking water consumed appropriate?

The identification of the unheated portion of the water seemed to have appropriate application to determining exposure to microbes. In the case of DBPs, NCEA decided to utilize total water consumption rather than a heated or unheated fraction (p.5-16). There are many reasons to question how this modifies the accuracy of the risk assessments for DBPs. Heating would clearly increase volatilization of some DBPs, but it may well accelerate the formation or degradation of others. Therefore, the assumption is justified, but the uncertainties should be noted.

d. Assess the validity of the assumption that the pathogen would not survive the pathways (referred to as a "preparation methods") from tap to consumption identified.

The question of "pathways" and "preparation methods" was difficult to identify specifically in the document. There was a discussion of the effect of heating water on pathogen survival, but that was all that could be identified. The assumption that *Cryptosporidium* would not survive in heated water may be correct. The extent of heating is a behavioral variable for different uses (especially for beverage preparation) that was not captured and how this affects the validity of this assumption was not discussed in the document.

e. Assess the validity and significance of the assumption that DBP concentrations do not change as a result of transport through the distribution system and many pathways through which water is consumed (e.g., boiling to prepare tea).

The assumption that DBPs do not change in distribution systems may be an acceptable assumption in this particular case study, but not an acceptable assumption generally. Making such an assumption is equivalent to assuming that free chlorine is always the primary disinfectant or that all source waters are the same.

f. Should other routes of exposure be included such as dermal and inhalation

routes? Would the contribution of these be significant?

There is no doubt that routes of exposure in addition to ingestion need to be accounted for in applications of the methodology to specific cases. However, the contribution of these routes need to be considered in the context of the integrated distribution of total exposure. If one adds the 95th percentile exposure via ingestion, to the 95th percentile of inhalation exposure to the 95th percentile exposure through the skin one will not arrive at the 95th percentile of total exposure. Clearly, the importance of various routes will vary depending upon the nature of the contaminant. Inhalation of aerosols has not been seriously considered, but as illustrated by outbreaks of Legionella several decades ago contaminants need not be volatile to present a significant inhalation hazard.

There is no doubt that under certain conditions dermal and inhalation exposure can make significant contributions to the overall exposure to disinfectants, their byproducts and microbial agents. Clearly some microbial agents in water are effectively spread through the inhalation route (i.e. Legionella), and certainly inhalation of DBPs during showering is well established. However, the method for treating such exposures needs careful consideration. EPA has indicated that in order to account for volatile chemicals, an extra 2 liters/day is considered (EPA, 1996). This value seems to be derived from studies that emphasized extremes in water usage rather than on how these various routes contribute to exposure in a population. If the Committee has an appropriate understanding of the issue, the question is an overall distribution of effective exposures to individual agents that need to be considered. Clearly, the effective dose will vary depending upon the physical-chemical properties of the byproduct and the relative effectiveness of inhalation versus ingestion as a mode of transmitting infections by different microbial agents. Therefore, it is not appropriate to use the approach suggested in the literature which treats inhalation and ingestion routes as some set fraction of overall exposure. The exposure estimates should be the population integral of these various routes with their associated confidence intervals. In any final analysis, these integrated exposures should include consideration of the pharmacokinetic variables of different routes of exposure to low doses of these chemicals in humans. Risks from DBPs should consider dermal and inhalation routes.

5. Health Conditions

a. Are the assignment and definition of health conditions appropriate for the progression of *Cryptosporidiosis*?

The DWC consolidated its comments on this point in section 2.b.ii of this Appendix.

b. Are the assignments and definition of health conditions for developmental, reproductive and cancer risks appropriate? Are latency periods and reversibility issues handled reasonably?

The DWC viewed the case study as an illustrative exercise, not a final analysis of the issue. Moreover, the results of the case study are largely trivial because of the overwhelming risks that are assigned to risk from *Cryptosporidium*. Nevertheless, there were methodological issues that should be addressed before the methodology is applied in a more definitive analysis.

There are significant difficulties in the way these hazards were defined and handled. Mostly, the draft document reflected a superficial treatment of the toxicological literature of most disinfectant byproduct classes and culminated in the assignment of toxicological properties to the rest of the total organic halogen present in the water across species and to lower doses. This is an unprecedented level of extrapolation beyond the available data. The diverse set of mechanisms (or modes of action) and the potential for sensitive populations to certain byproducts were not considered or identified as sources of uncertainty. Scattered throughout the Committee's comments are questions of whether some of the translation of information into human conditions are appropriate. These will not be repeated here. Of most concern in the context of this question is to make sure that it is clearly recognized that definition of health conditions at the human level based on animal data can raise serious concerns about the validity of the overall analysis when they are eventually converted into QALYs. Extreme care must be taken to ensure that conservatism on one side of the equation is not propagated in such a way that it distorts the final comparisons. A specific example of how this can become a problem is outlined in subsection c below.

For future case studies and further development of the method, the Agency will need to become considerably more invested in understanding the complex literature that is developing in this area than was apparent in this first attempt.

c. Are the uncertainties in this process identified?

Uncertainties in this process are not identified with sufficient specificity to drive research that could resolve these uncertainties. Several prior sections of this report address specific problems in this area and will not be repeated here (see Section 3 of this report).

A substantive problem was identified in the case study through discussions between Committee and National Center for Environmental Assessment (NCEA) personnel at the February, 1999 meeting that needs to be noted. Essentially, the problem was that data obtained in animal studies were used to characterize the risk "quantitatively" that is associated with disinfection byproducts. However, human data were used to estimate the resulting QALYs. This introduced a problem that was not explicitly identified in the analysis, in part because it was not apparent that this manipulation of the data was actually performed. The specific problem that could arise comes from the fact that the "risks" one estimates from the epidemiology data outstrip the toxicology data by 1-3 orders of magnitude, depending upon how the risks are actually calculated. Second, the epidemiology data introduce cancer sites that have not frequently been found in toxicological studies. This leads to a significantly broader band of uncertainty in the results of the analysis than is currently portrayed. These assumptions strike at the very heart of the analysis because they directly impact the QALYs assigned to the carcinogenic endpoints. In addition to the impact on the quantitation of the risks, it is well known that cancers at different sites carry significantly different implications for the quality of life as well as life expectancy once diagnosed.

This problem has other practical consequences in the case study. In all likelihood the reason this crept into the example was because there is no causal agent that can be associated with the bladder and colorectal cancer sites seen in humans. Consequently, there is no way to determine the impact of changing water treatment. However, if these risks are

real, they would completely drown out the risks that would be calculated from the DBPs that have been characterized in animal bioassays for cancer. The Committee recognizes the need to be pragmatic in these analyses, however, the analyst must be careful not to gloss over or “forget” about these shortcuts because the result could be an analysis that does not provide a true picture of the “state of the science”. Consequently, it is very important that these differences be captured in the uncertainty analysis. If this was the case, it was not made apparent in the document.

6. Common Health Metric

a. Is the use of Quality Adjusted Life Years (QALYs) appropriate?

The QALY concept seems to result in an appropriate common health metric for this case study. As indicated above, however, more attention must be paid to the string of assumptions that are inserted into deriving the QALY from experimental data. NCEA did not entertain alternatives. Now that they have had some experience with this metric, it might be appropriate to step back and assess alternatives.

In the context of water related problems, the Committee suggests that NCEA consider how acute outbreaks fit into this picture more specifically. If outbreaks are included in the analysis the number of “QALY’s” they effect may not be great enough to justify their inclusion in the analysis because they don’t occur that often. This would be a mistake. The morbidity and mortality associated with outbreaks is of greater impact than the morbidity and mortality resulting from estimates of endemic microbial infection and chronic disease from exposure to chemicals. This is because, while the morbidity and mortality in these epidemics is real and measurable, the endemic and chronic estimates are just that, “estimates”, extrapolations based on assumptions of exposure, dose response, and, in the case of chemical exposure, cross-species extrapolation. The proposed CRFM may not be suitable for directly comparing outcomes that are so different in their quality.

b. Is the assignment of QALYs given the health conditions appropriate?

The assignment of QALY costs to health outcomes in Chapters 5 and 6 required a great many assumptions, combining results from a number of different studies and adapting results derived in one situation to a quite different situation. For example, the study used the QALY cost of a year in severe pain derived in the Hamilton, Ontario study to estimate the QALY cost of a bout of severe acute illness, by dividing the one-year cost by 365 to get a one-day figure, and then multiplying by 14 for a two-week illness. These assumptions embedded in this procedure are obviously heroic. On the other hand, the authors really had little choice, for attempts to estimate QALYs directly for acute illness have not been very successful.

It would be useful to do a survey that directly pits against each other the two risks examined in the case study (acute GI disease and cancer after some period of latency). The survey results could be used to estimate directly the QALY cost of the various outcomes of these diseases, which could then be compared to the QALY costs estimated in the current study. A close correspondence between the survey-based QALY costs and those estimated in the case study would help build confidence in the method used to calculate QALYs. If there were no close correspondence, then survey results would nonetheless help analysts develop

better estimates of the QALY cost of acute nonfatal disease.

There does appear to be a problem with the estimate of the QALYs for the two alternative technologies for the AIDs subpopulation. The number of QALYs that the point-of-use filtration system would save appears to be about 22 QALYs per AIDS patient (11,636 QALYs for 429 persons), which is much greater than could reasonably be expected from the elimination of cryptosporidiosis in this population, given the short expected lifespan of these individuals (about two years, given the conditional survival probability of 0.53). As the document points out (p. 6-41) this QALY estimate is actually shared by a number of AIDS victims over the 20-year planning horizon; as patients die, others contract the disease and take their place. It still seems high on an individual basis, but impossible to say for sure without further information. Is the analysis supported by an explicit dynamic model of AIDS encompassing both survivorship and new cases, one that takes into account existing and anticipated future patterns of AIDS infectivity and survivability? Or is a simple “steady-state” assumption made, one that assumes that there are now X cases of AIDS and that in future years there will also be X cases? (If a steady-state is being assumed in the base case, it would seem that conditions would no longer be in steady-state after intervention, for AIDS patients would have a higher life expectancy while new AIDS cases would be added to the population at the same rate as the base case.) If there is a dynamic analytical model, it would be useful to include more information in the report. If there is no such model, it would be an important and useful addition the analysis. Considering the large share of the benefits that accrue to AIDS patients, it seems important to examine the dynamics of this population carefully.

APPENDIX B

Acronyms and Abbreviations

BCA	Benefit-Cost Analysis
CDW	Centrally Distributed Water
CE	Cost-Effectiveness
CRFM	Comparative Risk Framework Methodology
CSFII	Continuing Survey of Food Intake of Individuals
DBP	Disinfection Byproducts
DWC	Drinking Water Committee
IRIS	Integrated Risk Information System
MCLG	Maximum Contaminant Level Goal
MCL	Maximum Contaminant Level
MHI	Median Household Income
NAS	National Academy of Science
NCEA	National Center for Environmental Assessment
NODA	Notice of Data Availability
PDF	Probability Distribution Function
QALY	Quality Adjusted Life Year
QSAR	Quantitative Structure Activity Relationship
RO	Reverse Osmosis
SAB	U.S. EPA Science Advisory Board
SDWA	Safe Drinking Water Act Amendments of 1996
THM	Trihalomethanes
TOC	Total Organic Carbon
TOX	Unidentified Halogenated Fraction
WTP	Willingness to Pay

REFERENCES

- Ashby, J. and R.W. Tennant. (1994) Prediction of rodent carcinogenicity for 44 chemicals: Results. *Mutagenesis*, 9:7-15.
- Bull, R.J. and F.C. Kopfler. (1991) Health effects of disinfectants and disinfection byproducts. AWWA Research Foundation.
- Canada Ministry of Health and Welfare. (1981). *Tapwater Consumption in Canada*. Document No. 82-EHD-80. Public Affairs Directorate, Dept. of National Health and Welfare. Ottawa, Canada.
- Chappell, C. L., C.R. Sterling, and W. Jakubowski. (1998) *Cryptosporidium parvum* volunteer study: Infectivity, Illness, and Immunity. In Proceedings of the 1998 Drinking Water Progress Review Workshop for the 1995/1998 Science to Achieve Results (STAR) Grants. December 8-9, 1998. Arlington, VA. US EPA600/R-98/162. p. 16-17.
- Erschow, A.G., L.M. Brown, and K.P. Cantor. 1991. Intake of tapwater and total water by pregnant and lactating women. *Am. J. of Public Health*. 81:328-334.
- Helsel, D.R. (1990) "Less than obvious statistical treatment of data below the detection limit," *Environmental Science and Technology*. 24(12):1766-1774.
- Hertzberg, R.C., G. Rice, and L. Teuschler. (1998). Methods for health risk assessment of combustion mixtures. In *Hazardous Waste Incineration: Evaluating the Human Health and Environmental Risks*. S. Roberts, C. Teaf, and J. Bean, Eds. (In Press).
- Hurst, C.J., R. Clark, and S.E. Regli. (1996). Estimating the risk of acquiring infectious disease from ingestion of water. Ch. 4. In *Modeling Disease Transmission and Its Prevention by Disinfection*. C.J. Hurst, Ed. Cambridge Univ. Press. Cambridge, UK. P. 99-199.
- ILSI Working Group. (1996). A Conceptual framework to assess the risks of human disease following exposure to pathogens. ILSI Working Group; *Risk Analysis*, 16(6):841-47.
- Kodell, R.L., J.J. Chen, and D.W. Gaylor. (1995) Neurotoxicity modeling for risk assessment. *Reg. Toxicology and Pharmacol.* 22:21-29.
- LeChevallier, M.W., W.D. Norton, M. Abbaszadegan, and T.B. Atherholt. (1998). *Short term variability of Giardia cyst and Cryptosporidium oocyst concentrations in a surface water source used for potable water*. New Jersey Dept. of Envir. Protection, Div. of Science and Research. April.
- Morgenstern, R. (1998). *Regulatory Impact Analysis at the US Environmental Protection Agency*. Resources for the Future. Washington, DC.

- Morris, R.D., A-M. Audet, I.F. Angelillo, T.C. Chalmers, and F. Mosteller. (1992) Chlorination, chlorination byproducts, and cancer: A meta-analysis. *Am. J. of Public Health.* 82:955-963.
- National Research Council. (1983) *Risk Assessment in the Federal Government: Managing the Process.* National Academy of Science, Commission on Life Sciences, Committee on the Institutional Means for Assessment of Risks to Public Health. National Academy Press. Washington, DC. 191pp.
- Payment et al. (1991). Gastrointestinal health effects associated with the consumption of drinking water produced by point-of-use domestic reverse-osmosis filtration units. *Applied Environmental Micro.* 57:945-948.
- Payment, P., J. Siemiatycki, G. Richardson, G. Renaud, E. Franco, and M.A. Prevost. (1997). Prospective epidemiological study of gastrointestinal health effects due to the consumption of drinking water. *Int. J. Environ. Health Research.* 7:5-31.
- Pereira, M.A. and J.B. Phelps. (1996) Promotion by dichloroacetic acid and trichloroacetic acid of N-methyl-N-nitrosurea-initiated cancer in the liver of B6C3F1 mice. *Cancer Letters.* 102:133-141.
- Perz, J.F., F.K. Ennever, and S.M. LeBlancq. (1998) *Cryptosporidium* in tap water. Comparison of predicted risks with observed levels of disease. *Am. J. Epidemiology.* 147:289-301.
- Sobsey, M., A. Dufour, C. Gerba, M. LeChevallier and P.J. Payment (199) Using a conceptual framework for assessing risks to health from microbes in drinking water. *J. of the American Water Works Association.* 85:44-48.
- Stauber, A.J., R.J. Bull, and B.D. Thrall. (1998) Dichloroacetate and trichloroacetate promote clonal expansion of anchorage-independent hepatocytes *in vivo* and *in vitro*. *Toxicol. And Applied Phramacol.* 150:287-294.
- US Court of Appeals for the District of Columbia Circuit. (1999). May 14, 1999 opinion in the case of the American Trucking Associations, Inc., *et al.*, versus US Environmental Protection Agency (Opinion No.97-1440).
- US EPA Science Advisory Board. (1995). *Science Advisory Board (SAB) Review of Issues Related to the Regulation of Arsenic in Drinking Water.* US EPA SAB, Washington, DC. EPA-SAB-DWC-95-015.
- US EPA. (1997). *Exposure Factors Handbook. Vol. 1. General Factors.* Office of Research and Development. National Center for Environmental Assessment. EPA/600/P-95/002Fa. August.
- US EPA (1998). National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts. *Federal Register.* 63(241):69398.