American Water Works Association

The Authoritative Resource on Safe Water SM

Government Affairs Office
1300 Eye Street NW
Suite 701W
Washington, DC 20005
T 202.628.8303
F 202.628.2846
www.awwa.org

Headquarters Office
6666 W. Quincy Avenue
Denver CO 80235
T 303.794.7711
F 303.347.0804

June 11, 2007

Ms. Mabel Nichols
Office of Information and Regulatory Affairs
Office of Management and Budget
NEOB, Room 10201
725 17th Street
Washington, DC 20503

RE: Draft 2007 Report to Congress on the Costs and Benefits of Federal Regulations

Dear Ms. Nichols:

The American Water Works Association (AWWA) appreciates the opportunity to review and comment on the Draft 2007 Report to Congress on the Costs and Benefits of Federal Regulations as published on March 12, 2007 (72 FR 11061). AWWA has commented on several of the reports in the past, as Cost-Benefit Analysis (CBA) is an important tool for evaluating all rulemakings, but especially so for regulations issues under the Safe Drinking Water Act (SDWA). The 1996 SDWA Amendments elevated the importance of CBA by explicitly requiring consideration of costs and benefits in the development of drinking water standards.

The AWWA is an international, nonprofit, scientific and educational society dedicated to the improvement of drinking water quality and supply. Founded in 1881, the Association is the largest organization of water supply professionals in the world. Our 60,000 plus members represent the full spectrum of the drinking water community: treatment plant operators and managers, environmental advocates, engineers, scientists, academicians, and others who hold a genuine interest in water supply and public health. Our membership includes more than 4,700 utilities that supply roughly 80 percent of the nation's drinking water.

As previously mentioned, CBA is an important tool for SDWA rulemakings and the Office of Management and Budget (OMB) and the Environmental Protection Agency (EPA) need to ensure that the CBAs clearly reflect the risk-cost tradeoffs that drinking water regulations will impose on utilities' customers. Five out of the six regulations identified in the 2007 Draft Report as "major" rules costing more than $100 million per year are national drinking water regulations. In response to the majority of "major" rules being drinking water regulations, AWWA contracted with Dr. Robert Raucher of Stratus Consulting to conduct a detailed review of the 2007 Draft Report. s detailed review is

enclosed as part of these comments, along with previous critiques of the Economic Analyses (EAs) for the Stage 2 Disinfection By-Products Rule (DBPR) and the Long-Term 2 Enhanced Surface Water Treatment Rule (LT2ESWTR) as appendices to his review.

If you have any questions about these comments, please feel to call Alan Roberson or me in our Washington Office at 202-628-8303.

Yours Sincerely,

Thomas W. Curtis
Deputy Executive Director

cc: Ben Grumbles—USEPA OW
George Gray—USEPA ORD
Audrey Levine—USEPA ORD
Brian Mannix—USEPA OPEI
Cynthia Dougherty—USEPA OGWDW
Alan Roberson
Steve Via

FORMAL COMMENTS
BY THE
AMERICAN WATER WORKS ASSOCIATION
ON THE
DRAFT 2007 REPORT TO CONGRESS ON THE COSTS AND BENEFITS OF FEDERAL REGULATIONS, NOTICE AND REQUEST FOR COMMENTS

(March 12, 2007, 72 FR 11061)

EXECUTIVE SUMMARY

The American Water Works Association (AWWA) is pleased to submit this set of comments on the Office of Management and Budget (OMB) draft *2007 Report to Congress on the Costs and Benefits of Federal Regulations,* as printed in the *Federal Register* (Vol. 72, No. 47, March 12, 2007, p 11061). AWWA is dedicated to providing safe drinking water to the American public, and recognizes the importance of setting health-based standards that are balanced against the need to keep drinking water affordable.

Cost-Benefit Analysis (CBA) is an important tool for evaluating all rulemakings, but especially so for regulations issued under the Safe Drinking Water Act (SDWA). The 1996 SDWA Amendments have elevated the importance of CBA by explicitly requiring consideration of costs and benefits in the development of drinking water standards. Hence, the concerns raised here are not only about how benefits and costs are estimated, but also how they are compared to one another and interpreted in the standard setting context. Further, because the consumers who receive the benefits of drinking water standards are also the same group that will bear the costs, it is especially important that the CBAs clearly reflect the risk-cost tradeoffs that regulations will impose on them, as accurately as possible.

One of our main concerns with the draft *2007 Report to Congress* is that AWWA believes that the Environmental Protection Agency (EPA) is not giving adequate recognition of other related research when SDWA CBAs are developed. For example, AWWA typically reviews in detail each of the Economic Analyses (EAs) prepared for each SDWA rulemaking, and AWWA believes that our issues are not adequately addressed in the comment response documents or in later rulemakings. OMB relies solely on EPA-based estimates of the costs and benefits of the major drinking water rulemakings that it discusses. Relying solely on EPA-derived estimates provides a too-limited perspective of the likely benefits and costs of the regulations.

The discussion of the Stage 2 Disinfection Byproducts Rule (DBPR) is one notable example. There are many reasons that indicate that the EPA-based estimates of the benefits of this rule are considerably over estimated, and costs under estimated, and many of these reasons are detailed in the critique of the Stage 2 DBPR EA attached as Appendix A. As most drinking water regulations as considered “major” rules as defined by the federal government, OMB should provide a more in-depth and broadly

representative perspective on these regulations' likely levels of benefits and costs and not simply accept, report and, in essence, codify the EPA estimates.

In addition to AWWA's comments about specific rules and the associated estimates of their costs and benefits, AWWA also offers several insights and comments on how CBAs can and should be improved for evaluating and setting federal drinking water standards under the SDWA. More specifically, the Association offers insights and examples of ways to improve the quality, accuracy, and transparency of CBAs – and hence improve the effectiveness of the regulatory process and the standards promulgated – for federal rules issued under the SDWA. These comments are provided in response to OMB's request for "recommendations for improving the transparency, accountability, and effectiveness of the regulatory process, as well as for improving this Report" (p. 5). AWWA's observations for improving CBAs and associated regulatory processes include:

- Transparency, full disclosure, and replicability are essential. The methodologies utilized in developing a CBA and the intermediate (building block) results of the estimation process should be presented in a meaningful way, rather than only the final results. This will improve upon the current approach in which reviewers often are required to accept on faith the outcomes of a "black box" process, in which initial inputs may be known, and outputs described, but the intermediate steps in the CBA are not available for reality checks or public scrutiny.

- Central tendency "best" estimates need to be used. In addressing inevitable uncertainties underlying risk and other values in a CBA, extreme values should not be used for key inputs. "Best" estimates are not simply the average of the range. Instead, central estimates, along with plausible bounds and sensitivity analyses, are most suitable. Overly conservative assumptions (e.g., from the "precautionary principle" as applied in risk assessments) need to be identified and removed from CBAs.

- Benefits should be portrayed in the most useful and relevant metrics. For example, life years saved (LYS) should be estimated and used instead of (or alongside) estimates of premature fatalities.

- Sound economic principles should be applied in valuing benefits and costs. For example, where value of a statistical life (VSL) estimates are applied, appropriate latency and discounting should be used.

- Incremental benefits should be compared to incremental costs. This is an acceptable manner for assessing whether the regulatory choices may increase social welfare by maximizing net benefits.

- CBA results should be disaggregated to gauge important equity and efficiency considerations. For drinking water standards, key issues include the cost-benefit outcomes in small systems, and also cumulative regulatory impacts on affordability by low income households.

A. INTRODUCTION

The American Water Works Association (AWWA) is pleased to submit this set of comments on the Office of Management and Budget (OMB) draft *2007 Report to Congress on the Costs and Benefits of Federal Regulations,* as printed in the *Federal Register* (Vol. 72, No. 47, March 12, 2007, p 11061). OMB did an excellent job in assembling and reviewing the complex issues associated with analyzing the benefits and costs of the federal regulatory program. This is an important task, and a difficult one, and OMB is to be commended for its efforts.

AWWA is an international, nonprofit, scientific and educational society dedicated to the improvement of drinking water quality and supply. Founded in 1881, the Association is the largest organization of water supply professionals in the world. Our 60,000-plus members represent the full spectrum of the drinking water community: treatment plant operators and managers, environmental advocates, scientists, academicians, and others who hold a genuine interest in water supply and public health. Our membership includes more than 4,700 utilities that supply roughly 80 percent of the nation's drinking water. Given the depth and breadth of AWWA's representation, these comments should be taken as reflecting the predominant view of the nation's drinking water professionals. It is therefore appropriate that these AWWA comments be heard on behalf of the drinking water community in general.

AWWA is principally interested in providing safe drinking water to the American public, and recognizes the importance of health-based standards that are balanced against the need to keep drinking water affordable. AWWA's primary focus is thus on regulations issued by the U.S. Environmental Protection Agency (EPA), under the 1996 Safe Drinking Water Act (SDWA) Amendments, that establish Maximum Contaminant Levels (MCLs) and treatment requirements for chemical and microbial contaminants in drinking water. We also are interested in the broader issue of evaluating health policies and environmental protection choices that span the wide spectrum of federal programs that affect water quality and public health.

AWWA believes that judicious use of cost-benefit analysis (CBA) has an important role to play in helping society ensure that investments in drinking water quality (and other programs) yield the greatest public health protection benefits possible. CBA can help identify which contaminants are the highest priorities to regulate, and the most suitable level at which to regulate them.

CBA is an important tool for evaluating all rulemakings, but especially so for regulations issued under the SDWA. The 1996 Amendments have elevated the importance of CBA by explicitly requiring consideration of costs and benefits in the development of drinking water standards. Hence, the concerns raised here are not only about how benefits and costs are estimated, but also how they are compared to one another and interpreted in the standard setting context. Further, because the people who receive the benefits of drinking water standards are also the same group that will bear the costs, it is especially important

that the CBAs clearly and accurately reflect the risk-cost tradeoffs that regulations will impose on them.

For many years, AWWA has been carefully reviewing CBAs for national primary drinking water regulations issued by EPA under the Safe Drinking Water Act (SDWA). We have extensively commented on many significant cost-benefit issues in our lengthy comments on EPA's proposals for radon, radionuclides, arsenic, the groundwater rule, and the group of rules known as the Microbial/Disinfection By-Product (M/DBP) Cluster. We have also taken a look backwards at the CBAs in the final drinking water regulations. We were an active participant in the 2001 review of the arsenic regulation, and still have some unresolved concerns with the EPA CBA and the Agency's related documentation for the arsenic rulemaking.

As part of developing comments on EPA's proposed rules, the drinking water community as a whole has invested thousands of member person-hours and spent millions of dollars with the hope of improving the regulatory development process. EPA has made some improvements in the quality of its CBAs for drinking water regulations. However, despite considerable efforts by Association staff, members, and experts on AWWA's behalf, and some improvement from EPA, significant concerns remain about the quality and interpretation of many of the CBAs developed by EPA for drinking water regulations.

Benefits and costs must be estimated by the federal agencies and presented to the public in a manner that is as objective, accurate, transparent, and replicable as possible. Section 1412(b)(3)(C) of the SDWA details the requirements of the health risk reduction and cost analyses that must be conducted. Sections 1412(b)(5) and 1412(b)(6) detail additional health risk and cost considerations. Further, policy decisions must be made on suitable and objective interpretations of how the benefits compare to costs, including instances where significant uncertainties exist and/or important benefits or costs cannot be readily quantified within the CBA framework.

Accordingly, AWWA applauds OMB's efforts to ensure that CBAs are done properly and consistently by federal regulatory agencies. Further, AWWA believes OMB's efforts in its annual *Report to Congress* are important, and has comments to offer on the issues raised in the 2007 draft. The Association's comments are organized by "chapter" as presented in the OMB draft.

B. ESTIMATES OF THE BENEFITS AND COSTS OF FEDERAL REGULATIONS (CHAPTER I)

Benefits and Costs of "Major" Rules: Stage 2 Disinfection Byproducts Rule

The OMB draft *2007 Report to Congress* provides a summary of the benefits and costs of the seven "Major Rules" it reviewed in federal fiscal year 2006. One of these major rules is the National Primary Drinking Water Regulation: Stage 2 Disinfection Byproducts Rule (referred to below as the "Stage 2" or "DBP" rule). AWWA has several

observations regarding the EPA benefit and cost estimates used by OMB for this major Stage 2 DBP rule, and most of these points are provided later in these comments.

For the Stage 2 DBP Rule, OMB reports a range of annual benefits of approximately $600 million to $1.5 billion, and annualized costs of $74 million to $76 million (in year 2001 U.S. dollars, from Table 1-4, p. 12). AWWA believes that there is considerable evidence that the EPA estimated benefits are overstated and costs understated – perhaps to a significant degree. For example, EPA's epidemiologically based estimates of bladder cancer are inconsistent with its interpretation of the toxicological evidence, and the difference is implied risk is more than a factor of 400 using the toxicological evidence. And this is just one of the many factors that could significantly lower the benefits for this specific rulemaking.

The reasons why the Association believes that there are significant inaccuracies in the Stage 2 DBP Rule benefits and costs as reported by OMB (based on EPA's estimates) are discussed in greater detail at the end of these comments, and are important in their own right. However, the more critical issue for AWWA is that by reporting only the EPA-based CBA results, OMB lowers the value and credibility of its own *Report to Congress*. OMB needs to continue to pressure EPA to address the many limitations in the EPA analyses that OMB itself has often recognized and tried to get EPA to address.

AWWA believes that instead of simply reporting EPA and other federal agency CBA results, OMB also should describe the key limitations, uncertainties, and potential inaccuracies in EPA (and other agency) CBA outcomes. OMB also should provide alternative estimates of benefits and costs, when they have been developed in credible fashion by OMB and/or other parties. OMB can do this by using reliable information as is often available from public comments and other sources, and from OMB's own reviews of EPA's CBAs (which are part of OMB's routine activities in its oversight of major federal rulemaking activities).

As there were only seven major federal rules for FY 2006 on which OMB needs to report, this should not be an unreasonable burden (especially as OMB already is deeply involved in reviewing and critiquing the agency CBAs during the rulemaking process). The value added by diving more deeply into the basis of the estimated benefits and costs, and providing readers with alternative and equally (or more) valid estimates would be a useful public service and help lead to a better regulatory process.

Finally, AWWA would like to encourage OMB to report not just the total benefits and costs of the rules evaluated, but also the incremental benefits and costs. Where feasible, it is quite informative to show how the benefits and costs change as one moves from one regulatory option to the next more stringent alternative.

Impact of Federal Regulations on State, Local, and Tribal Governments

Six federal regulations issued over the past 10 years are identified in OMB's draft *2007 Report to Congress* as having estimated annual costs of $100 million or more that are

imposed on state, local, or tribal governments. Five out of these six are National Primary Drinking Water Regulations (and the sixth is a storm water discharge rule that will tend to impact the same communities, and in many cases the very same utilities as affected by the drinking water standards).

The federal drinking water standards noted for their impact on state and local communities are:

- Stage 1 DBP Rule (1998)
- Interim Enhanced Surface Water Treatment Rule (IESWTR) (1998)
- Arsenic Rule (2001)
- Long Term 2 Enhanced Surface Water Treatment Rule (LT2) (2005)
- Stage 2 DBP Rule (2006)

The EPA-derived cost estimates for these rules combined amounts to more than $1.5 billion per year (in year 2001 dollars). This is the sum of the EPA-based costs for each of the above rules, as reported by OMB to be approximately $700 million (Stage 1) , $400 million (IESWTR), $200 million (Arsenic), $130 million (LT2), and $70 million (Stage 2) (pp. 24, 25).

The impact of these rules deserves greater consideration and discussion than is provided in the draft *2007 Report to Congress*. First, the cumulative cost impact is significant – over $1.5 billion for 5 rules issued within an 8-year period. Second, the true cumulative impact is likely to be considerably higher than the EPA-based estimate. This is because: (1) AWWA strongly believes most of the EPA-derived cost estimates are considerably understated, and (2) there are other federal drinking water rules that have (or are about to) also impact these same communities (e.g., covering radionuclides, unregulated contaminant monitoring, and groundwater). AWWA encourages both OMB and EPA to allocate sufficient budgetary resources to conduct an appropriate retrospective study on the cost of drinking water regulations. AWWA is considering contracting for a retrospective cost of compliance study for the arsenic regulation, as any new treatment installed for arsenic compliance is unique compared to conventional treatment so that appropriate cost allocation would be relatively simple. To place this in perspective, consider that the total national cost of SDWA regulations promulgated since 1986 has been estimated to be about $5 billion per year (Raucher and Cromwell, 2004).

It also is important to recognize that these cumulative costs fall overwhelmingly on the local and regional water supply utilities that are owned and run by cities, towns, counties, and other localities. The fiscal impacts of these federal drinking water regulations are not broadly dispersed throughout the U.S. economy. Instead, they are concentrated on the households that bear the higher water rates needed to pay for compliance with these standards. Federal fiscal assistance, such as through the Drinking Water State Revolving Fund (DWSRF), is often touted as the solution, but the annual DWSRF appropriation is only 0.3% of the 20-year need based on EPA's Drinking Water Needs Survey (Roberson, 2006). If the current level of the DWSRF were maintained for the next 20 years, the total appropriations would be approximately 6% of the $276.8 billion 20-year funding need.

Additionally, the DWSRF is primarily a loan program (even at low- or no-interest rates), and these loans have to paid back through rate increases. This leaves most of the cost burden on localities and the households who reside in them. Given the other cost-impacting issues faced by these water utilities (e.g., infrastructure renewal, enhanced security measures, securing additional source waters), there are significant fiscal pressures on America's water utilities.

In addition, although the OMB draft *2007 Report to Congress* clearly identifies these rules and provides a brief description of their costs and anticipated benefits, there is no discussion of the *impacts* these collective rules impose on communities. It would be useful for OMB to assess and report the level and types of *impacts* these rules have on communities, rather than to simply acknowledge that these rules exist and impose costs.

One important type of impact to address is the potential for health risk tradeoffs faced by the low income and fixed income households bearing the regulatory compliance costs of drinking water regulations. As noted above, the total national cost of Safe Drinking Water Act (SDWA) regulations promulgated since 1986 has been estimated to be about $5 billion per year (Raucher and Cromwell, 2004). CBAs conducted in development of these standards assert that health benefits are at least equal to these expenditures. By comparison, if there are 10 million low income households facing an average water and sewer bill of more than $400 per year (Rubin, 2005), this equates to more than $4 billion per year to be raised from low-income households. To whatever extent that portion of water revenue is obtained at the expense of other health-related expenditures by these households, the resulting detriment to health could rival the magnitude of improvements to health intended under the SDWA.

The National Rural Water Association recently released a White Paper: *The Relationship Between Household Financial Distress and Health: Implications for Drinking Water Regulation* (Rubin, Raucher, and Harrod, 2007)[1]. The 2005 Behavioral Risk Factor Surveillance System (BRFSS) used in this research indicates a strong correlation between financial distress and several illnesses and other adverse health outcomes such as diabetes and cardiovascular disease. Water utilities and regulators need to remain mindful that public health is their core business, and that there is conceivably as much health impact at stake in the manner in which they obtain revenue from low income households as there is in treating the water to higher standards.

Finally, while AWWA does not have specific insights to provide on the CBA results described for the storm water rulemaking, the Association does wish to raise the important but oftentimes overlooked connection between actions taken under the Clean Water Act (CWA) and the 1996 SDWA Amendments (or on drinking water quality issues in general). There are many regulatory actions that have an impact on source water quality, and AWWA believes source water protection benefits should be fully recognized when wastewater or land use regulations are being evaluated. Rules governing pesticides or other agricultural chemicals also may provide important drinking water benefits. AWWA has commented previously on several pesticide reregistrations

[1] Available at http://www.nrwa.org/whitepapers/2007/Rubin%20tradeoff%20health%20risk%20final.pdf

by recommending that manufacturers be required to conduct more sampling in drinking water sources as part of the reregistration process. In general, it often makes sense to control the source of pollution rather than its consequences, and source water protection also is more consistent with the "user pays" principle.

C. TRENDS IN BENEFIT AND COST ESTIMATES (CHAPTER II)

AWWA does not have specific comments about the information provided in this chapter of the draft *2007 Report to Congress*. However, the Association suggests that it would be instructive in this Report (or future year Reports) to examine the *cumulative* costs of federal rulemakings on key sectors of the economy over time. For example, the cumulative impact of federal drinking water standards on local water utilities and the households they serve is significant, and important to recognize on both the cost and benefit sides of the ledger.

D. UPDATE ON IMPLEMENTATION OF THE INFORMATION QUALITY ACT (CHAPTER III)

AWWA does not have specific comments about the information provided in this chapter of the draft *2007 Report to Congress*. AWWA has followed the implementation of the Information Quality Act and internally discussed filing a petition on the Stage 2 DBP rule on the fetal loss estimate, but has not been involved in any petition at this time.

E. AGENCY COMPLIANCE WITH THE UNFUNDED MANDATES ACT (CHAPTER IV)

This chapter of the draft *2007 Report to Congress* focuses on regulations issued in federal FY 2006, and describes the Stage 2 DBP drinking water rule as one of the federal regulations that will have a sizable fiscal impact on local government entities. The draft *2007 Report to Congress* also notes that although this rulemaking is an unfunded federal mandate, some impacted communities may be eligible for federally supported financial assistance through the Drinking Water State Revolving Fund.

In reference to the information provided in this chapter of the *2007 Report to Congress*, AWWA would like to reiterate a few points raised earlier in these comments. Specifically, the Association believes it is important to point out that (1) the fiscal impact of the Stage 2 DBP rule will likely be significantly higher than indicated by the EPA-based cost estimate; and (2) it is important to recognize the cumulative fiscal impact of the federal drinking water regulatory program over time, rather than focus on the latest fiscal year alone. AWWA recommends that in future EAs on future rulemakings, EPA present a cumulative cost of all its drinking water regulations.

F. RECOMMENDATIONS FOR IMPROVING THE REGULATORY PROCESS: IMPROVING COST-BENEFIT ESTIMATES OF FEDERAL REGULATIONS

Introduction

The draft *2007 Report to Congress* includes a request from OMB for "recommendations for improving the transparency, accountability, and effectiveness of the regulatory process, as well as for improving this Report" (p. 5). Below, the Association offers insights and examples of ways to improve the quality, accuracy, and transparency of CBAs – and hence improve the effectiveness of the regulatory process and the standards promulgated – for federal rules issued under the SDWA.

AWWA understands the difficulties and frustrations of trying to evaluate federal agency CBAs for national regulations. Below, the Association offers insights and examples based on its review of federal rules issued under the SDWA, and offers ideas to facilitate OMB's and other reviews in the future. The role of CBA is important in all rulemakings, but especially now under the SDWA, where CBA has a role in standard setting. Hence, the concerns raised here are not only about how benefits and costs are estimated, but also how they are interpreted in the standard setting context.

AWWA Recommendations for Improving EPA's Drinking Water CBAs

For many years, AWWA has been carefully reviewing CBAs for federal rulemakings issued by EPA under the SDWA. Considerable efforts by Association staff, members, and experts on AWWA's behalf, and OMB reviewers have led to some limited improvements in some recent EPA EAs for drinking water standards. OMB Circular A-4, providing guidance on best practices for regulatory analysis, has also been useful in this regard. Nonetheless, there remain concerns about many of the key CBAs developed by EPA on drinking water issues.

EPA's drinking water CBAs often have been difficult to review or replicate, and/or appear to be in error in several respects. In certain respects, EPA's CBAs also have not conformed to the explicit requirements of the SDWA (notably, CBA-related provisions under various portions of section 1412). A discussion of issues related to EPA's CBA for the Stage 2 DBP Rule and the Long Term 2 Enhanced Surface Water Treatment (LT2) Rule are provided later in this review, as well as in the attached appendices (Appendix A and B are AWWA-supported reviews of EPA's EAs for the Stage 2 DBP and the LT2 microbial rules, respectively).

The key issues that tend to re-occur in EPA's CBAs for drinking water rules include:

- Lack of transparency, replicability, and consistency. In several instances, it has been difficult or impossible to follow what the Agency has done in its analyses. Key citations are not always made available (or refer back to other documents such as internal Agency memoranda, until the trail ends short of the key facts). Results from intermediate steps are not always provided, so it is impossible to "put the pieces

together" to determine the source of numerical discrepancies. This means that in certain instances the public must accept the EPA estimates on faith. This is at odds with sound practice, and also does not conform to the SDWA requirement for public information [section 1412(b)(3)(B)].

There also has sometimes been a lack of consistency among studies in terms of data, methods, or assumptions applied. Inconsistency would not be a problem if the changes over time reflected a steady evolution toward improved methods and data. Regrettably, this is not the case for some of OGWDW's CBAs.

- Reliance on overly conservative assumptions and default values when estimating benefits. In the face of uncertainty, risk assessors traditionally apply the "precautionary principle" in determining what exposure levels are "safe." This is done through use of uncertainty factors, reliance on upper confidence limits and a linear dose-response model for carcinogens, and the application of other practices that are intentionally designed to avoid understating risk.

  The use of the precautionary principle is perhaps suitable in defining an unenforceable, risk-free, or *de minimus*, goal such as an MCLG. For other purposes, however, it is inappropriate for risk assessment to include such conservative policy judgments. For its CBAs, EPA should provide unbiased estimates of risk that are in turn suitable for risk *management* applications such as the use of CBA in standard setting . Otherwise, the risk assessments will lead to a considerable overstatement of likely benefits.

  Benefits analyses need to reflect "best estimates" (or suitable probability distributions) for key exposure, dose-response, latency period, and benefits valuation issues. This is not only sound economics and policy analysis, but it also is required under the Section 1412(b)(3)(B) of the SDWA.

- Not providing or considering meaningful *incremental* comparisons of benefits to costs. EPA has typically used a comparison of total benefits to total costs in evaluating MCL options, even though the SDWA (section 1412) and economic principles dictate that incremental benefits should be compared to incremental costs to maximize net social benefits. A comparison of total benefits and costs indicates only whether or not a rule is a break-even proposition, and this is an insufficient basis for choosing whether or not to regulate, or how stringently to set the standard. This is especially important for SDWA regulations because many of the standards (MCLs) have been set where the total costs exceed the total benefits if the central tendencies were appropriate examined or if benefits were appropriately discounted.

- Reluctance to use meaningful measures of risk reduction benefits, such as "Life Years Saved" (LYS). Reduced risks of premature fatalities need to be viewed in the context of the amount of increased longevity (years of life extension) provided by a regulation. This provides a more meaningful way to interpret regulations, some of

which may reduce premature fatalities early in life, and others that are aimed more at risks faced late in life. EPA has steadfastly adhered to the more generic, less informative "lives saved" approach, even though EPA (in its own Clean Air Act analysis) and other agencies (e.g., FDA) have published more informative CBAs using the LYS approach.

EPA has refused to estimate LYS in drinking water regulations because the Science Advisory Board (SAB) raised some concerns with valuing LYS on the basis of adjusting estimates of the Value of a Statistical Life (VSL). Nonetheless, even if there are concerns about developing a monetary estimate of the value of a statistical life year (VSLY), this is no basis for refusing to at least quantify the degree of life extension provided by regulatory options developed under the SDWA regulatory program.

Finally, Quality Adjusted Life Years (QALYs) and related measures may also be worth considering as alternative ways of portraying the health-related risk reduction benefits of drinking water and other federal regulations.

- Lack of more systematic approaches for considering unquantified benefits and costs within CBA and standard setting. In some instances, important benefits or costs may not be readily quantified or portrayed in dollar value terms. In these instances, the unquantified or omitted benefits and costs need to be suitably considered in the regulatory decision-making process -- they should neither be ignored nor given undue weight. EPA's CBAs for drinking water standards have sometimes failed to use available information on unquantified outcomes in an informative manner, despite examples being provided to EPA in our comments on several past rulemakings.

- Inadequate consideration of the affordability of rulemakings. EPA focuses only on median household incomes, and does not adequately consider the cumulative impact of multiple pending regulations on household water bills. This is a particular concern when considering low income households and residents of smaller communities.

- Masking significant regional economic impacts under a national context. Several SDWA regulations have regionalized impacts due to contaminant occurrence being concentrated in a few geographic areas (e.g., uranium, radium). The regional impact of these rules can be significant, but this important perspective is masked when the Agency uses only a national aggregate analysis which makes the overall impact seem modest.

All of above recommendations (and more) are also part of the recommendations in one of the following four recent reports on drinking water regulatory actions:

- *Report to Congress: Small Systems Arsenic Implementation Issues* (March 2002)

- *Drinking Water: Revisions to EPA's Cost Analysis for the Radon Rule Would Improve Its Credibility and Usefulness* (GAO, February 2002)

- *Report of the Arsenic Cost Workgroup to the National Drinking Water Advisory Council* (August 2001)

- *Arsenic Rule Benefits Analysis: An SAB Review* (August 2001)

While the recommendations from these reports (and other reports dating back several years) have been known and well articulated for several years, EPA needs to fully incorporate these recommendations in its drinking water CBAs.

EPA took some small steps in addressing these recommendations in the proposed Stage 2 Disinfection By-Products Rule and the Long-Term 2 Enhanced Surface Water Treatment Rule.[2] However, EPA still fell short in incorporating these recommendations in the proposed CBAs performed for the Stage 2 DBP Rule and LT2 Rule (as detailed later in these comments).

Additional Recommendations for Improving the Regulatory Process

OMB should develop a* cumulative *programmatic assessment of the federal drinking water regulatory program. To supplement the more limited rule-by-rule and year-by-year approach in the annual *Reports to Congress*, AWWA recommends OMB develop a comprehensive evaluation of drinking water regulatory programs and their associated benefits and costs. This will improve the general comprehension of the tradeoffs and priorities that can be made within the program to ensure the greatest public health protection is obtained at the lowest possible cost.

AWWA also encourages OMB to evaluate other public health and safety programs, so that suitable and informative cross-program comparisons can be made. For example, it is important to grasp how investments in traffic safety, occupational health, food and drug programs, air quality, wastewater and nonpoint source controls, and source water protection compare to those in drinking water in terms of how the benefits compare to the costs. This could be done in an informative manner similar to that presented in Graham et al (1998), which facilitates cross-program comparisons of regulatory results and efficiencies based on a common cost-effectiveness metric (i.e., cost per life year saved).[3]

[2] The regulatory structure for these rules was approved through a lengthy Federal Advisory Committee (FACA) process. Therefore, the CBAs did not have any impact on the options selected for these specific standards. Nonetheless, it is important to ensure that the CBAs are of the highest quality possible.

[3] Graham, J.D., P.S. Corso, J.M. Morris, M. Segui-Gomez, and M.C. Weinstein. 1998. *Evaluating the Cost Effectiveness of Clinical and Public Health Measures.* Published in the *Annual Review of Public Health*, Volume 19. J.E. Fielding, editor.

OMB should consider credible CBA estimates provided by parties other than those developed by the federal agencies responsible for the rulemaking. Many times, CBA-relevant insights are available from nonfederal parties. OMB should consider CBA-relevant information regardless of its source, and include the data and results of CBAs performed by independent organizations (unless there is clear reason to believe the analyses lack technical credibility).

Thus, OMB should review and use analyses and results provided by stakeholders and other outside parties, where those results are credible. OMB should not rely exclusively on EPA or other federal agency data and results. OMB should consider input from other informed parties, and use the input data and/or findings of these other parties where such stakeholder analyses are objective, credible, and reliable.

Greater credence should be provided to public comments, and related technical supporting materials. OMB should review Public Comments submitted in response to proposed rulemakings, and also review Agency "Comment-Response Documents" to better ensure that EPA (and other federal regulatory agencies) fully and suitably considers public comments. AWWA, other stakeholders, and researchers face frustration when their public comments and technical reports are virtually ignored by regulating agencies. AWWA has often submitted relevant data and credible analyses to EPA within the rulemaking process, and frequently the Association feels that the materials are largely ignored by EPA. OMB may find it informative to review the submitted comments, as well as Agency "Comment-Response Documents" to determine if the comments are suitably considered (rather than being noted only in pro-forma manner).

Technical Issues with EPA's CBA Estimates for the Stage 2 DBP Rule

Research and analyses sponsored by AWWA, the AWWA Research Foundation (AwwaRF), and/or other stakeholders have often found significant differences between their data and results and those provided by EPA. Because the draft *2007 Report to Congress* relies solely on EPA's estimates of the benefits and costs of the Stage 2 Disinfectants and Disinfection Byproducts Rule (as one of seven major regulations issued by the federal government in 2006), this section provides a summary of the critique of the EPA's Economic Analysis (EA) of the proposed rule and the full critique is attached as Appendix A. This is intended to highlight specific reasons why AWWA believes the EPA-reported estimates are seriously in error and should not be the only values shown in the OMB *2007 Report to Congress*.

AWWA's review of the EA methods, data, results, and presentation by EPA reveals some good efforts by the Agency, but also several critical areas of concern that generate erroneous and misleading results and that require extensive improvement. Our major observations and findings with respect to the EA are as follows:

1. Overall, EPA has overstated the net benefits of the proposed rule.

a. EPA has significantly understated how many water systems might be affected by the rule as proposed, and the per system cost of compliance may also be underestimated based on EPA's assessment of the "increment" between Stages 1 and 2. Therefore, EPA probably has underestimated the total regulatory costs and impacts considerably.

b. EPA, through its selective and questionable use of epidemiological evidence, has significantly overstated the risks associated with DBP exposures, and thus the Agency overestimates the per system benefits of the proposed rule to a considerable degree.[4]

2. The occurrence analysis provides a potential underestimate of the number of utilities (and distribution site locations) with compliance issues for the Stage 2 rule. This is due to the inherent and well-recognized limitations of the SWAT model (and the underlying ICR data) for the purposes of estimating plant-specific DBP profiles, especially for values near the tail of the distribution (i.e., away from the mean).

3. EPA's exposure assessment contributes to considerably overestimated levels of bladder cancer risks associated with elevated DBP levels.

a. For EPA's risk estimates based on epidemiological evidence, the implicit exposure assessment applied by EPA is that Americans of the regulation-impacted future (2013 and beyond) will have levels and durations of chlorinated surface water ingestion identical (or nearly so) to the tap water exposure patterns observed over the many past decades of the epidemiological study populations. Given increased use of bottled water and in-home treatment devices, as well as increased residential mobility, the levels and durations of exposure are likely to be far lower for the regulation-impacted population than for the epidemiological study groups. This implies that, based on ingestion-related exposure issues alone, any epidemiologically associated risks and benefits applied to the proposed Stage 2 rule would be lower than those estimated by EPA.

b. For EPA's risk estimates based on toxicological evidence, the explicit exposure assessment is based on 2 liters of Community water System (CWS) tap water consumption per day over 70 year exposure duration. These are standard assumptions used in conservative risk assessments, but are extremely overstated values for a realistic assessment of exposures for a benefit-cost analysis. As a result of the tap water ingestion factor alone, EPA overstates the exposure (and hence risks) posed by DBPs by a factor of 2 or 3 or more, compared to risk estimates derived when EPA's own estimates of central tendency ingestion rates for CWS waters, and

[4]. The extent of overstatement for national level benefits is mitigated to some extent by EPA's under-prediction of systems affected. However, the nation level net benefit (benefits – costs) is still distorted.

increased bottled water use, are properly considered.

4. EPA's use of epidemiological evidence to develop quantified estimates of bladder cancer risk (and risk reduction benefits attributable to the rule as proposed) is considerably flawed and ill-advised.

   a. A suitable association between bladder cancer and DBPs is lacking. EPA makes a fundamental and unsubstantiated leap of faith by taking evidence associating bladder cancer with long-term exposure to chlorinated surface waters and equating that to a basis for estimating risks posed by THM4, HAA5, and/or other DBPs as may coincidently be reduced through the proposed Stage 2 rule. EPA clearly recognized this limitation in its Stage 1 rulemaking with a lower bound of the cases avoided of zero. However, in the Stage 2 rulemaking, this lower bound disappeared and the Agency moved beyond the bounds of sound science while still using the same information.

   b. EPA applies an internally inconsistent and biologically implausible set of assumptions by implicitly applying a linear dose-response function for DBPs (as consistent with carcinogens that are "initiators") and at the same time imposing a very short cessation lag for risk reduction (as consistent with carcinogens that act as "promoters" rather than initiators, and only if the promoting agents are evident at effective doses).

   c. EPA's epidemiologically based risk estimates are wholly inconsistent with its interpretation of the toxicological data. The difference in implied risk is perhaps more than a factor of 400 lower using the toxicological evidence (Crawford-Brown, 2003). The wide divergence suggests that the epidemiological studies are compromised by one or more significant confounding factors associated with exposure to chlorinated surface water.

5. The Agency's development of fetal loss estimates stretches the limits of sound science, imprudently makes the Agency appear to be fear-mongering, and represents a probable violation of the federal data quality guidelines.

   a. The epidemiological evidence does not support quantified analysis. The FACA Advisory Committee reviewed this body of research, and after two years of deliberation and consultations with the world's leading experts, concluded that the existing research lacked sufficient strength and evidence of causality to support any quantified risk assessment. EPA has acted contrary to the Committee's directive, and any semblance of sound science, in developing its "illustrative calculation" of fetal loss.

   b. EPA's review, presentation, and use of the epidemiological evidence is biased and misleading. The Agency has engaged in a "data dredging" exercise in which it downplays or ignores inconclusive, inconsistent, or

contrary results, and instead highlights and applies only the observed positive associations between DBPs and reproductive or developmental effects. The Agency has simply provided a laundry list of positive associations noted amongst the many endpoints and studies. This is not in keeping with the Agency's own *Carcinogenicity Guidelines,* nor is it consistent with EPA's higher standards of assessment typically applied by the Agency to epidemiological evidence.

c. EPA, and the underlying studies that EPA selects for application, do not account for other probable causative agents and exposure pathways. Therefore, there is a likelihood of significant confounding that is not considered.

d. EPA's quantitative "illustration" not only is based on a biased and inappropriate interpretation of the underlying epidemiological evidence, but also is significantly flawed in how it interprets the ICR data to estimate cases avoided due to Stage 2. EPA misinterprets the ICR data to estimate a 69% reduction due to Stage 2 in fetal loss associated with chlorinated drinking water. In contrast, a more appropriate evaluation of the same data (such as focusing on the site-specific ICR data, and on brominated species instead of chloroform) suggests that a 1% reduction is more likely.

6. EPA's valuation of nonfatal bladder cancer cases is problematic.

a. The high end estimate is based on a proportional relationship to the "Value of Statistical Life" (VSL), based on a "benefits transfer" of results drawn from a study by Magat et al. (1996). EPA does not develop its benefits transfer in accordance with its own *Guidelines for Preparing Economic Analyses*, and important documentation needed to assess the reliability and applicability of the Magat et al. study has not been made available for review (despite considerable efforts to obtain the important information from study authors, their university, and the Agency office that funded the original research). Our own assessment suggests that the EPA estimates appear overstated.

b. EPA's use of data on the willingness to pay (WTP) to avoid chronic bronchitis is also questionable. Again, the Agency has not provided any evidence of a careful and systematic benefits transfer (as would be consistent with the Agency's own *Guidelines*). Our perspective is that the differences between the health endpoint studied and nonfatal bladder cancer suggest that the literature-based estimate may be too high for application to the Stage 2 rulemaking context.

7. The Agency should develop and portray estimates of life years saved (in addition to its estimates of numbers of cancers avoided). Further, the cost per life year

saved should be provided for public review, decision-maker deliberations, and stakeholder consideration.

8. EPA's estimation of the cancer risk reduction cessation lag is highly questionable, and probably overstates how quickly bladder cancer risks might be reduced by a reduction in drinking water DBP exposures.

    a. The cessation lag concept applies to a portion of the population, whereas latency applies to others. EPA needs to consider how to reflect both concepts in its temporal analysis of risk reductions.

    b. EPA directly applies data from reduced lung cancer risks among smokers who cease their tobacco use in a direct transfer to bladder cancer risk reductions from lowering levels of DBPs in drinking water. The numerous and profound problems with this direct transfer of the smoking cessation results to the Stage 2 context include issues of different agents, biokinetics, cell turnover rates, bioaccumulation, distinctions between initiators and promoters (and the levels at which either or both occur), exposure pathways, target organs, and cessation versus marginal exposure reduction.

    c. Given the vast uncertainties associated with the cessation lag estimation, and the numerous problems in applying the tobacco results directly to the DBP context without adjustment, EPA should have, at a minimum, conducted sensitivity analyses to reveal how different the key EA results would be if alternative assumptions were applied. EPA also should have provided much more systematic peer review of this issue.[5]

9. Various portions of the EPA EA may well be in violation of EPA and OMB information quality guidelines. The fetal loss estimates may be the most serious and apparent transgression of these guidelines.

    a. EPA itself states that it "does not believe the available evidence provides an adequate basis for quantifying potential reproductive/developmental risks." Yet despite this assessment, the Agency nonetheless proceeds to develop empirical estimates. EPA characterizes these empirical estimates as "illustrative calculations" yet the Agency highlights their numeric results throughout the rulemaking package and EA.

    b. The EPA fetal loss analysis is not objective either in its presentation or in terms of its substance. EPA's presentation of the epidemiological evidence is neither even-handed nor unbiased, and its illustrative calculation is not accurate, reliable, or unbiased. EPA cherry-picked what data and studies

[5] In its EA for the final Stage 2 rulemaking, EPA did offer a limited sensitivity analysis of two additional cessation lag scenarios. However, these alternatives also overlook the more fundamental scientific and peer review issues and, therefore, are not particularly useful.

to use and ignored a more balanced approach that would have more fully captured the mixed and inconsistent results evident in the body of relevant research.

c. EPA does not appear to have conducted any substantive expert peer review of its fetal loss approach or results, nor did the Agency conduct a pre-dissemination review (as clearly required under the OMB guidelines for information quality).

d. The Agency's interpretation of the cessation lag estimates from tobacco smoking and lung cancer, and its application of these findings to the DBP context and bladder cancer, may reflect another area in which the spirit and letter of information quality guidelines were both ignored.

10. EPA estimates the costs of the proposed Stage 2 rule as a simple increment over Stage 1. At the time of the FACA negotiations, this was a legitimate approach because Stages 1 and 2 were anticipated to be implemented within a relatively short time of each other, such that the compliance efforts would be highly integrated. However, as the time between Stage 1 and Stage 2 has grown, the selection and implementation of Stage 1 compliance strategies have become more separate from the upcoming Stage 2 requirements, and higher costs are now likely to be incurred for Stage 2 than EPA projects under its incremental costing approach.

11. EPA's presentation of regulatory costs, affordability, and benefits is overly aggregated, and fails to reveal how affordability, costs, and benefits vary across system size categories.

12. EPA fails to provide any incremental net benefits analysis. Instead, the Agency simply provides national aggregate estimates of total benefits and total costs. The need for an incremental perspective has been clearly articulated in previous SAB and NDWAC reviews, yet the Agency fails to provide this simple yet highly informative portrayal of its findings.

These and other points are discussed in greater detail in the Stratus Consulting critique of the Stage 2 DBP rule EA, which is attached as Appendix A.

Technical Issues with EPA's CBA Estimates for the LT2 Rule

The EPA CBA for the proposed LT2 Rule had numerous shortfalls. The Economic Analysis (EA) and associated support documentation offered extensive detail and information. However, EPA should have found a better balance in the support documentation so that it provided not only complete, but also the most critical information to interested and involved parties. To find this balance, EPA should use more

fundamental, informative, and simple analyses of core components rather than using more sophisticated approaches for some less important aspects of the EA.

In some critical elements of the EA, the agency made powerful assumptions that could have significant impacts on the final results of the EA. The agency did not always clearly articulate what assumptions are being made and often presented a one-sided view of relevant uncertainties and data limitations to derive its interpretation. In some instances where the agency made key assumptions, the supporting analysis lacked sensitivity analyses based on equally or more plausible alternative assumptions.

Our major observations and findings with respect to the LT2 Rule EA included:

1. Overall, we believed EPA considerably overstated the occurrence and risks associated with endemic levels of cryptosporidium in finished waters, and thus the agency overstated the benefits of the proposed rule to a considerable degree.

2. The ICRSS data indicated a much smaller percentage of systems will end up in bins 3 and 4 under the proposed rule than do the analyses based on the ICR data, implying that the net benefits (benefits minus costs) of the proposed rule may be 20% of the high end estimates shown by EPA (all else equal). The ICRSS data would be better predictors than the ICR data of what the impact of the rule would be as proposed.[6]

3. EPA applied a Bayesian interpretation to the ICR and ICRSS data that was suspect and driven by unsubstantiated and perhaps extreme assumptions. For example, EPA imposed an assumption that only 1 out of every 1000 "zeroes" observed in the database is truly a zero. The agency estimated occurrence and risk based on a presumption that 999 out of every 1000 observed zeroes in the database were instead, one oocyst or more.

4. EPA's exposure assessment was based on considerably over-estimated levels of direct ingestion of CWS-provided waters. Relevant exposures (and, hence, risks) may be overstated by a factor of 2 or 3 when direct ingestion rates for CWS waters, and increased bottled water use, are properly considered.

5. The infectivity dose-response relationship applied by EPA was subject to considerable uncertainty and probably overstated the risk associated with exposures to an infectious oocyst by a significant degree.

    a. The underlying clinical studies used extremely high doses relative to oocyst levels in finished waters (levels of oocysts ingested of 23,000 to 2.3 billion times higher than now found in finished waters) and relied on

[6] The ICRSS data are more indicative of what the rule's impacts will be because they (1) probably are more accurate than the ICR data (ICRSS results are based on Method 1622/1623 with higher recovery rates than the IFA method applied in the ICR data) and (2) reflect the method (1622/1623) that utilities will apply in their compliance monitoring.

extremely small number of subjects and strains (between 14 and 29 subjects, for each of only 3 strains).

b. The results of the clinical studies were interpreted liberally, based on a "presumed infection" approach that assumed that any subject with symptoms had cryptosporidiosis, even when several of the symptomatic subjects had no documented infection (e.g., via positive oocyst shedding). EPA's risk estimates were overstated to the extent that reported symptoms could be attributable to causes other than cryptosporidiosis.

c. The results of the clinical studies were interpreted via complex statistical models that were driven by -- and highly sensitive to -- unsubstantiated assumptions. While the modeling approaches used by EPA in the EA were suggested by the SAB, the obscurity of the presentation and the sensitivity of the results to the model assumptions (e.g., increasing a key estimated mean risk parameter by a factor of 4 or 5 over the level found in the peer reviewed published literature) revealed the need for more transparency, continued scientific discourse, and greater use of sensitivity analyses in portraying the possible risk levels.

6. The extent by which EPA's risk model overstated risks can be viewed, in part, by comparing the agency's estimated number of waterborne cases of cryptosporidiosis at the pre-LT2 baseline to its estimated reduction in cases due to the proposed LT2 rule:

a. EPA estimated the pre-LT2 baseline (i.e., post IESWTR) to be between 60,000 and 111,000 cases per year.

b. The agency's risk model used for the LT2 rule benefit-cost analysis predicted 256,000 to over 1,000,000 cases per year will be avoided due to the rule as proposed.

c. Therefore, EPA estimated a reduction in cases that is up to 9+ times higher than the number of cases it stated existed at the baseline.

7. EPA should explore the soundness and implications of its questionable assumption that the risk of illness (as well as severity and duration of illness) was independent of dose. The morbidity assessment -- used to project the number, severity, and duration of illnesses due to a possible infection – was based exclusively on results from the Milwaukee outbreak of 1993, where oocyst levels were much higher, exposure durations much longer, and opportunities for secondary spread and exposure more pervasive than anticipated under the endemic low dose exposure context addressed by the proposed rule.

8. EPA's use of an "enhanced" cost of illness (COI) approach to value avoided cases of nonfatal cryptosporidiosis was highly problematic. The approach was a

significant departure from standard economics practice, did not appear to have been subjected to expert peer review, and yielded results that seem implausible and unrealistic compared to other well-established risk valuation benchmarks.

9. EPA's presentation of regulatory costs and benefits was overly aggregated, and failed to reveal how affordability and net benefits vary across system size categories or across other relevant program elements in the proposed rule (e.g., reservoir covering, filtered versus unfiltered systems).

These and other points are discussed in greater detail the Stratus Consulting critique of the LT2 rule EA, which is attached as Appendix B.

Appendix A.
Critique of the Economic Analysis for the Proposed Stage 2 Disinfectants and Disinfection Byproduct Rule, Final Report, Stratus Consulting

Stratus Consulting

Critique of the Economic Analysis for the Proposed Stage 2 Disinfectants and Disinfection Byproducts Rule

Final Report

Prepared for:

AWWA
1401 New York Ave., NW
Suite 640
Washington, DC 20005

Prepared by:

Robert Raucher, PhD
David Mills
Lauraine Chestnut
Stratus Consulting Inc.
PO Box 4059
Boulder, CO 80306-4059
(303) 381-8000

January 14, 2004
SC10365

Critique of the Economic Analysis for the Proposed Stage 2 Disinfectants and Disinfection Byproducts Rule

Final Report

Prepared for:

AWWA
1401 New York Ave., NW
Suite 640
Washington, DC 20005

Prepared by:

Robert Raucher, PhD
David Mills
Lauraine Chestnut
Stratus Consulting Inc.
PO Box 4059
Boulder, CO 80306-4059
(303) 381-8000

January 14, 2004
SC10365

Contents

Executive Summary

This report provides a critique of the U.S. Environmental Protection Agency (EPA) Economic Analysis (EA) of the proposed Stage 2 Disinfectants and Disinfection Byproducts Rule (referred to hereafter as the Stage 2 or DBP rule). The EA provides a benefit-cost analysis of the proposed rule. Our review of the EA methods, data, results, and presentation by EPA reveals some good efforts by the Agency, but also several critical areas of concern that generate erroneous and misleading results and that require extensive improvement.

Our major observations and findings with respect to the EA are as follows:

1. Overall, EPA has overstated the net benefits of the proposed rule.

    a. EPA has significantly understated how many water systems might be affected by the rule as proposed, and the per system cost of compliance may also be underestimated based on EPA's assessment of the "increment" between Stages 1 and 2. Therefore, EPA probably has underestimated the total regulatory costs and impacts considerably.

    b. EPA, through its selective and questionable use of epidemiological evidence, has significantly overstated the risks associated with DBP exposures, and thus the Agency overestimates the per system benefits of the proposed rule to a considerable degree.[1]

2. The occurrence analysis provides a potentially significant underestimate of the number of utilities (and distribution site locations) with compliance issues for the Stage 2 rule. This is due to the inherent and well-recognized limitations of the SWAT model (and the underlying ICR data) for the purposes of estimating plant-specific DBP profiles, especially for values near the tail of the distribution (i.e., away from the mean).

3. EPA's exposure assessment contributes to considerably overestimated levels of bladder cancer risks associated with elevated DBP levels.

    a. For EPA's risk estimates based on epidemiological evidence, the implicit exposure assessment applied by EPA is that Americans of the regulation-impacted future (2013 and beyond) will have levels and durations of chlorinated surface

1. The extent of overstatement for national level benefits is mitigated to some extent by EPA's under-prediction of systems affected.

water ingestion identical (or nearly so) to the tap water exposure patterns observed over the many past decades of the epidemiological study populations. Given increased use of bottled water and in-home treatment devices, as well as increased residential mobility, the levels and durations of exposure are likely to be far lower for the regulation-impacted population than for the epidemiological study groups. This implies that, based on ingestion-related exposure issues alone, any epidemiologically associated risks and benefits applied to the proposed Stage 2 rule would be lower than those estimated by EPA.

b. For EPA's risk estimates based on toxicological evidence, the explicit exposure assessment is based on 2 liters of CWS tap water consumption per day over a 70 year exposure duration. These are standard assumptions used in conservative risk assessments, but are extremely overstated values for a realistic assessment of exposures for a benefit-cost analysis. As a result of the tap water ingestion factor alone, EPA overstates the exposure (and hence risks) posed by DBPs by a factor of 2 or 3 or more, compared to risk estimates derived when EPA's own estimates of central tendency ingestion rates for CWS waters, and increased bottled water use, are properly considered.

4. EPA's use of epidemiological evidence to develop quantified estimates of bladder cancer risk (and risk reduction benefits attributable to the rule as proposed) is considerably flawed and ill-advised.

a. A suitable association between bladder cancer and DBPs is lacking. EPA makes a fundamental and unsubstantiated leap of faith by taking evidence associating bladder cancer with long-term exposure to chlorinated surface waters and equating that to a basis for estimating risks posed by THM4, HAA5, and/or other DBPs as may coincidently be reduced through the proposed Stage 2 rule. EPA itself recognized this limitation in its Stage 1 rulemaking, but for Stage 2 the Agency has moved beyond the bounds of sound science in how it uses the same information.

b. EPA applies an internally inconsistent and biologically implausible set of assumptions by implicitly applying a linear dose-response function for DBPs (as consistent with initiators) and at the same time imposing a very short cessation lag for risk reduction (as consistent with promoters rather than initiators, and only if the promoting agents are evident at effective doses).

c. EPA's epidemiologically based risk estimates are wholly inconsistent with its interpretation of the toxicological data. The difference in implied risk is perhaps more than a factor of 400 lower using the toxicological evidence (Crawford-

Brown, 2003). The wide divergence suggests that the epidemiological studies are compromised by one or more significant confounding factors associated with exposure to chlorinated surface water.

5. The Agency's development of fetal loss estimates is completely at odds with sound science, imprudently makes the Agency appear to be fear-mongering, and represents a probable violation of the federal data quality guidelines.

   a. The epidemiological evidence does not support quantified analysis. The FACA Advisory Committee reviewed this body of research, and after two years of deliberation and consultations with the world's leading experts, concluded that the existing research lacked sufficient strength and evidence of causality to support any quantified risk assessment. EPA has acted contrary to the Committee's directive, and any semblance of sound science, in developing its "illustrative calculation" of fetal loss.

   b. EPA's review, presentation, and use of the epidemiological evidence is biased and misleading. The Agency has engaged in a "data dredging" exercise in which it downplays or ignores inconclusive, inconsistent, or contrary results, and instead highlights and applies only the observed positive associations between DBPs and reproductive or developmental effects. The Agency has simply provided a laundry list of positive associations noted amongst the many endpoints and studies. This is not in keeping with the Agency's own *Carcinogenicity Guidelines,* nor is it consistent with EPA's higher standards of assessment typically applied by the Agency to epidemiological evidence.

   c. EPA, and the underlying studies that EPA selects for application, do not account for other probable causative agents and exposure pathways. Therefore, there is a likelihood of significant confounding that is not considered.

   d. EPA's quantitative "illustration" not only is based on a biased and inappropriate interpretation of the underlying epidemiological evidence, but also is significantly flawed in how it interprets the ICR data to estimate cases avoided due to Stage 2. EPA misinterprets the ICR data to estimate a 69% reduction due to Stage 2 in fetal loss associated with chlorinated drinking water. In contrast, a more appropriate evaluation of the same data (such as focusing on the site-specific ICR data, and on brominated species instead of chloroform) suggests that a 1% reduction is more likely.

6. EPA's valuation of nonfatal bladder cancer cases is problematic.

    a. The high end estimate is based on a proportional relationship to the "Value of Statistical Life" (VSL), based on a "benefits transfer" of results drawn from a study by Magat et al. (1996). EPA does not develop its benefits transfer in accordance with its own *Guidelines for Preparing Economic Analyses*, and important documentation needed to assess the reliability and applicability of the Magat et al. study has not been made available for review (despite considerable efforts to obtain the important information from study authors, their university, and the Agency office that funded the original research). Our own assessment suggests that the EPA estimates appear overstated.

    b. EPA's use of data on the willingness to pay (WTP) to avoid chronic bronchitis is also questionable. Again, the Agency has not provided any evidence of a careful and systematic benefits transfer (as would be consistent with the Agency's own *Guidelines*). Our perspective is that the differences between the health endpoint studied and nonfatal bladder cancer suggest that the literature-based estimate may be too high for application to the Stage 2 rulemaking context.

7. The Agency should develop and portray estimates of life years saved (in addition to its estimates of numbers of cancers avoided). Further, the cost per life year saved should be provided for public review, decision-maker deliberations, and stakeholder consideration.

8. EPA's estimation of the cancer risk reduction cessation lag is highly questionable, and probably overstates how quickly bladder cancer risks might be reduced by a reduction in drinking water DBP exposures.

    a. The cessation lag concept applies to a portion of the population, whereas latency applies to others. EPA needs to consider how to reflect both concepts in its temporal analysis of risk reductions.

    b. EPA directly applies data from reduced lung cancer risks among smokers who cease their tobacco use in a direct transfer to bladder cancer risk reductions from lowering levels of DBPs in drinking water. The numerous and profound problems with this direct transfer of the smoking cessation results to the Stage 2 context include issues of different agents, biokinetics, cell turnover rates, bioaccumulation, distinctions between initiators and promoters (and the levels at which either or both occur), exposure pathways, target organs, and cessation versus marginal exposure reduction.

c. Given the vast uncertainties associated with the cessation lag estimation, and the numerous problems in applying the tobacco results directly to the DBP context without adjustment, EPA should have, at a minimum, conducted sensitivity analyses to reveal how different the key EA results would be if alternative assumptions were applied. EPA also should have provided much more systematic peer review of this issue.

9. Various portions of the EPA EA may well be in violation of EPA and OMB information quality guidelines. The fetal loss estimates may be the most serious and apparent transgression of these guidelines.

   a. EPA itself states that it "does not believe the available evidence provides an adequate basis for quantifying potential reproductive/developmental risks." Yet despite this assessment, the Agency nonetheless proceeds to develop empirical estimates. EPA characterizes these empirical estimates as "illustrative calculations" yet the Agency highlights their numeric results throughout the rulemaking package and EA.

   b. The EPA fetal loss analysis is not objective either in its presentation or in terms of its substance. EPA's presentation of the epidemiological evidence is not even-handed or unbiased, and its illustrative calculation is not accurate, reliable, or unbiased. EPA cherry-picked what data and studies to use and ignored a more balanced approach that would have more fully captured the mixed and inconsistent results evident in the body of relevant research.

   c. EPA does not appear to have conducted any substantive expert peer review of its fetal loss approach or results, nor did the Agency conduct a pre-dissemination review (as clearly required under the OMB guidelines for information quality).

   d. The Agency's interpretation of the cessation lag estimates from tobacco smoking and lung cancer, and its application of these findings to the DBP context and bladder cancer, may reflect another area in which the spirit and letter of information quality guidelines were both ignored.

10. EPA estimates the costs of the proposed Stage 2 rule as a simple increment over Stage 1. At the time of the FACA negotiations, this was a legitimate approach because Stages 1 and 2 were anticipated to be implemented within a relatively short time of each other, such that the compliance efforts would be highly integrated. However, as the time between Stage 1 and Stage 2 has grown, the selection and implementation of Stage 1 compliance strategies have become more separate from the upcoming Stage 2

requirements, and higher costs are now likely to be incurred for Stage 2 than EPA projects under its incremental costing approach.

11. EPA's presentation of regulatory costs, affordability, and benefits is overly aggregated, and fails to reveal how affordability, costs, and benefits vary across system size categories.

12. EPA fails to provide any incremental net benefits analysis. Instead, the Agency simply provides national aggregate estimates of total benefits and total costs. The need for an incremental perspective has been clearly articulated in previous SAB and NDWAC reviews, yet the Agency fails to provide this simple yet highly informative portrayal of its findings.

1. Introduction

This report provides a review and critique of the Economic Analysis (EA) (U.S. EPA, 2003a), developed by the U.S. Environmental Protection Agency (EPA), in support of its proposed rulemaking for the Stage 2 Disinfectants and Disinfection Byproducts (DBP) Rule (referred to hereafter as the Stage 2 or DBP rule), published in the *Federal Register* on August 18, 2003 (U.S. EPA, 2003b). In essence, the EA is a benefit-cost analysis, and as such it contains many critical elements that support the rationale for the rulemaking and that are used to predict the rule's expected impacts on public health and water costs. The EA includes issues of occurrence, exposure, dose-response relationships, quantified risk characterization, and the valuation of health risk reductions. Issues related to compliance technologies, and their performance and costs, are also integral to the EA.

1.1 Objectives

The objective of this critique is to focus on key aspects of the Agency's empirical analysis, and highlight where there appear to be actual or potential errors or omissions of consequence in the EA. Our primary concerns are with the quality, transparency, replicability, objectivity, clarity, and reliability of the data and methods used to develop and present estimates of the benefits and costs of the proposed rule. Our focus is on the proper use and interpretation of sound methods and data, the presentation of the approach and outcomes in a clear and unbiased manner, and the drawing of suitable inferences and policy interpretations given the uncertainties and other limitations inherent in the data and analyses. The overall objective is to help EPA identify and remedy potential errors and limitations in its own EA, and thereby help EPA generate more reliable, accurate, and policy-useful benefit-cost analyses for the final Stage 2 and other rules in the future.

No attempts are made here to develop or report independent estimates of the proposed rule's benefits, costs, or incremental net benefits. Instead, we focus on identifying aspects of EPA's analysis that appear to require correction or, at a minimum, further investigation and documentation. The overall goal is to help identify areas in which improvements can and should be made, so that EPA's EA of the final Stage 2 rule — and of other future rulemakings — can embody improved quality, integrity, and meaning for the important public health protection policies the Agency pursues in accordance with the mandates of the Safe Drinking Water Act (SDWA). We hope and expect that in the EA supporting the final Stage 2 rule, EPA will embody the changes and improvements noted in this report so that the Agency can provide the public and decision-makers with a more accurate and informative assessment of benefits and costs.

1.2 Overview of Critique

Our critique of the EPA EA and associated Agency-provided documentation focuses on the need to find a better balance and sense of prioritization in the EA's analyses and presentation. For example:

- EPA is to be commended for provided a vast quantity of detail and back-up documentation. On the other hand, there is too much material to effectively review within the comment period. More important, EPA is inconsistent in that its documentation is sometimes lacking in terms of not providing enough information on some core issues, while offering mountains of data on other matters of sometimes lesser significance. In general, more balance and prioritization is needed in terms of detail and documentation.

- EPA is to be commended for attempting to quantify benefits to the greatest degree feasible, so that benefits can be compared to costs. On the other hand, EPA has in places stretched credibility and the bounds of "good science" in its efforts to generate estimates. "Good science" must always be the guiding principle, and empirical estimates based on interpretations or assumptions that have strayed from that principle are potentially misleading and a disservice to the public, stakeholders, and decision-makers.

- EPA has taken some good strides forward by providing considerable discussion and sophisticated numeric evaluation for several of the uncertainties and variabilities (e.g., using Monte Carlo simulations) for some aspects of the analysis. On the other hand, EPA neglects to detail, justify, or fully explore some of the most fundamental of its assumptions. In the face of these core uncertainties, sensitivity analyses are essential for evaluating the impact of core assumptions at key junctures of the analysis. EPA needs to find a better balance by using more fundamental, informative, and simple analysis of core components rather than using more sophisticated approaches for some lesser aspects of its analysis.

Hence, at the core of our critique is the message that EPA needs to take better stock of its analyses, determine what components are most critical in terms of driving the benefit or cost estimates, and focus its attention (and that of the reviewers) on those aspects. Models, analytic tools, and documentation should be presented in a way that sheds light rather than obfuscates and overwhelms attempts at good faith public review.

1.3 Outline of this EA Critique

Following this introductory chapter, the following information is provided:

- Chapter 2 provides a schematic overview of the EPA benefits analysis for the proposed Stage 2 rule, along with a summary of the key concerns identified by our review.
- Chapter 3 explores issues with the first set of steps in the analysis related to predicted changes in the quality of distributed water, in terms of both reduced temporal averages and reduced "peaks" that may occur at some times in some locations within a distribution system.
- Chapter 4 explores exposure issues and their impacts on the analysis.
- Chapter 5 discusses the dose-response elements of the analysis, notably the potential association of exposures to elevated DBPs with bladder cancers and adverse reproductive effects.
- Chapter 6 critiques the valuation of risk reductions, notably the monetized values assigned to nonfatal cancers, and the manner in which the timing (e.g., cessation lags) for bladder cancer cases are addressed by the Agency.
- Chapter 7 evaluates the Agency's cost analyses and the comparison of benefits to costs.

1.4 Conclusion

The EA and associated EPA-provided documents offer extensive detail, information, and background material. It is evident that the EPA has explored many aspects of the analysis and has provided a considerable body of documentation. However, some key elements of EPA's work are not sufficiently documented or detailed for an effective review, and this has hampered our evaluation. Overall, the Agency needs to find a more suitable balance.

In some very critical elements of the EA, the Agency makes strong assumptions that can have a significant impact on the final results, yet the Agency does not always clearly articulate what assumptions it is making or it appears to take a one-sided view of the uncertainties and data limitations to derive its interpretation. More important, in these instances of strong assumptions, EPA has in several instances failed to offer sensitivity analyses based on equally or more plausible alternative assumptions. EPA should follow the basic tenants of its own *Guidelines for Economic Analyses* and (a) be explicit regarding its core assumptions, (b) document the basis for those assumptions, and (c) develop some useful sensitivity analyses to evaluate and convey the impact of these core uncertainties on the outcomes of the analysis.

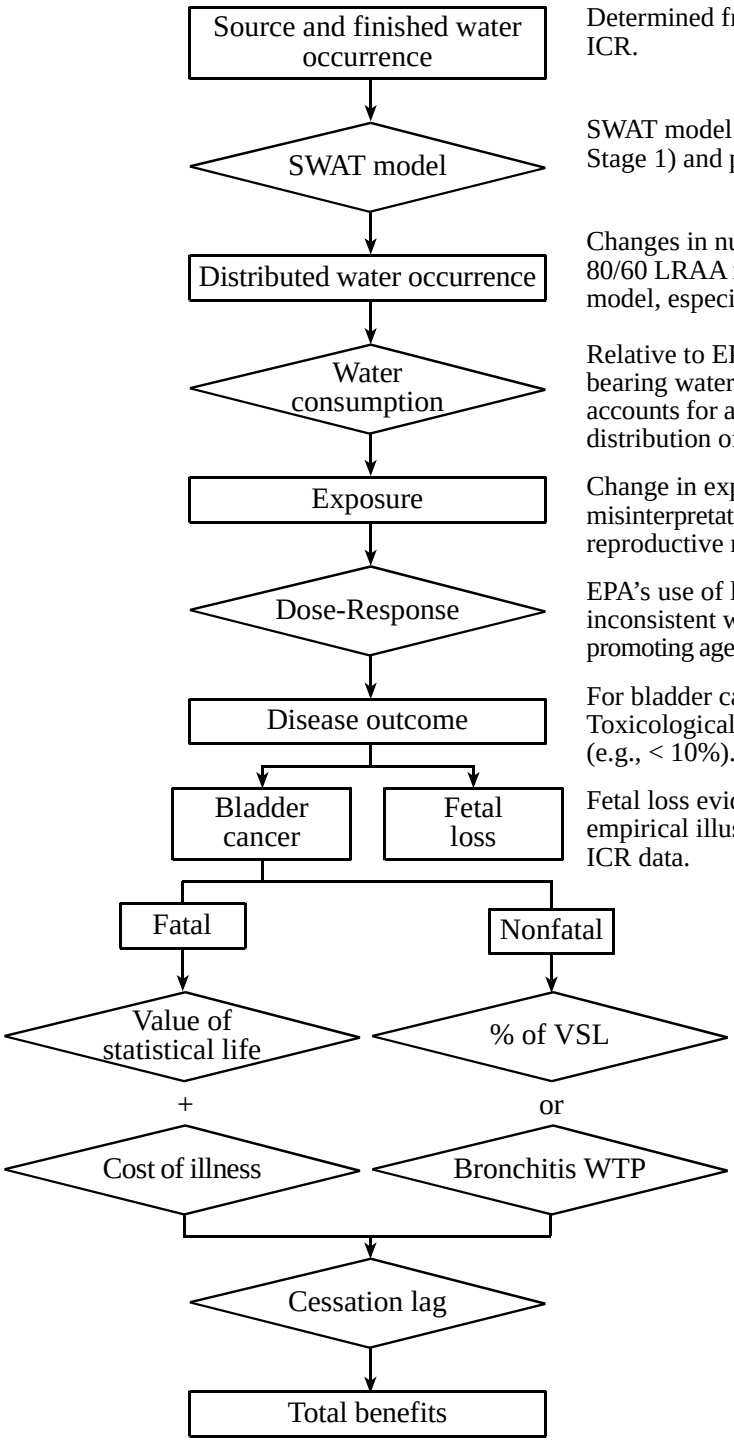
2. Overview of Benefits Analysis

To best understand the nature of the EPA analysis, a simplified flow diagram is shown in Exhibit 2.1 to identify and describe the key steps of the benefits assessment. The same underlying logic and data also underlie the basis for the cost analysis (although the path of the cost analysis naturally differs from that of the benefits after the initial occurrence assessment).

Along with each critical step in the analysis, Exhibit 2.1 displays a summary of some of the key issues and concerns regarding each of these steps. Additional detail is then provided in subsequent chapters of this report.

It is apparent when examining the issues identified in Exhibit 2.1 that there are many ways in which the risk reduction benefits developed by EPA may be overstated, all else equal. The "big picture" view of the EPA analysis, as afforded by Exhibit 2.1, allows us to glean how much the overall EPA results might be altered if alternative (but equally or more plausible) assumptions and data interpretations were investigated. Compounding the changes at each step as depicted in the figure, the overall estimate of benefits derived by EPA could be an order of magnitude or more larger than a more plausible and likely estimate. Detailed discussions are provided in the chapters that follow.

Other aspects of EPA's approach may lead to an underestimation of how many water systems may be affected by the Stage 2 rule as proposed. This implies that EPA may have significantly underestimated costs (and, if more systems are affected, the benefits may also be understated, all else equal). Overall, our impression is that EPA has understated the impact of the proposed rule in terms of costs. While some higher-than-projected effects may also yield more widespread DBP reduction efforts than projected by EPA, on net, benefits may still be overstated.



Determined from DBP monitoring data from NRWA and ICR.

SWAT model projections used to develop baseline (post Stage 1) and post Stage 2 DBP levels.

Changes in number, duration, and levels of "peaks" above 80/60 LRAA may be mis-estimated due to nature of SWAT model, especially relative to IDSE sites.

Relative to EPA's estimates, the mean ingestion of risk-bearing water may decline by 30% to 50% or more if one accounts for averting behavior and uses CWS-based USDA distribution of tap water ingestion.

Change in exposure, as derived by EPA from its misinterpretation of the ICR, overstates EPA's "illustrative" reproductive risk numbers.

EPA's use of linear dose-response function for DBPs is inconsistent with Agency's cessation lag premised on promoting agents (for which nonlinear dose-response applies).

For bladder cancer, epidemiological evidence misapplied. Toxicological evidence suggests far fewer cases avoided (e.g., < 10%).

Fetal loss evidence not suitable for quantification, plus empirical illustration based on erroneous interpretation of ICR data.

EPA's estimate of WTP to reduce risk of nonfatal bladder cancer, based on % of VSL, requires greater scrutiny, and key documentation for underlying study by Magat et al. is unavailable for review.

Cessation lag based on tobacco smoking and lung cancer probably is inappropriate and too short for DBPs and bladder cancer.

Total benefits may be overstated considerably, based on numerous issues noted above.

Exhibit 2.1. Benefits analysis process.

3. DBP Occurrence

The occurrence analysis of DBP levels is a important component of the benefit-cost analysis. It is used to project the Stage 2 baseline by estimating how many systems have DBP levels above regulatory thresholds of interest (i.e., MCL options) post Stage 1. The occurrence analysis is also used to project how much those DBP levels will decrease because of compliance with a given regulatory option. Because the preferred regulatory alternative for Stage 2 entails a locational running annual average (LRAA) that must be complied with based on water quality at various representative elevated DBP sites within the distribution system, projecting baseline and post-compliance DBP levels (and their composition) is a very challenging exercise.

The rule as proposed is aimed at providing more uniform DBP levels throughout a distribution system. More specifically, it is intended to minimize the possibility that DBP levels in some locations within a distribution network could be far higher than the 80 µg/L and 60 µg/L running annual average (RAA) for THM4 and HAA5 that were established for utilities (plant-wide rather than locational) under Stage 1. Therefore, Stage 2 as proposed will reduce locational variability in DBPs levels through impacted distribution systems by changing DBP levels temporally (e.g., seasonally) or spatially within distribution network, or both.

- At specific "relatively high DBP" sites within distribution systems that exceed the LRAAs at baseline, average levels of HAA5 and THM4 will be reduced by compliance. In the process of reducing site-specific LRAAs, it is likely that the temporal variation at that site will be reduced as well (e.g., that the relatively high DBP values that may occur from time to time will be reduced, thus shaving some of the "peak" DBP values).

- It is also possible that some of the compliance strategies will reduce average and relatively high temporal DBP levels at other locations throughout the distribution system in addition to the specific regulation-driven sites with representatively high DBP levels.

Thus, while compliance with the rule is geared specifically at select sites and running averages at those sites, it is also likely that in many instances the "peaks" also will be shaved at those sites, and that averages and "peaks" may be reduced at other locations as well.

EPA's approaches to address the complex occurrence aspects outlined above are based on efforts made to inform deliberations by the FACA Advisory Committee. There are several well recognized limitations with the data, methods, and results derived from the approaches developed during the FACA process. These limitations affect the reliability and interpretation of the occurrence analysis in terms of projected costs and benefits.

Many of the significant limitations of the occurrence analysis reflect the inherent limitations of the SWAT model (in conjunction with the ICR data that provide essential DBP level inputs) to predict baseline and post-Stage 2 DBP levels, especially for extreme events that deviate from national-level means. The SWAT/ICR model/data combination was not designed for and does not predict well DBP results on a plant-specific basis (although the FACA Committee accepted the outcomes as a basis for considering national means and standard deviations), nor does it predict well values at the tails of the distribution (it is intended for use at or near the means). As a consequence, the EPA occurrence projections probably underestimate occurrence and hence compliance costs (and risk reductions, all else equal).

The SWAT- and ICR-related limitations, and their implications, include potential significant underprediction of Stage 2 compliance for reasons including:

- The IDSE-based sites at which utilities will need to monitor and come into compliance with the proposed LRAAs in many cases will have higher DBP levels than levels measured at the ICR monitoring locations. This is because the IDSE will be used to identify sites with the highest anticipated DBP levels (whereas even the ICR "MAX" site probably does not capture the worst case locations for DBPs, although it is intended to reflect a high-end residence time).

- The proposed regulatory compliance monitoring requirements — requiring monitoring in the hottest month and then every 90 days thereafter — are more likely to reveal exceedences of the proposed MCLs than projected in the EA using the ICR-based estimates derived by SWAT. In part, this is due to capturing higher temperatures in the compliance monitoring regime as proposed (compared to temperatures most likely represented in the ICR monitoring data).

- Predictions of "peak" levels based on SWAT are especially suspect, and practitioners who helped develop and apply the model believe a significant underprediction has resulted. The model is limited to single hit projections and thus misses variability within and across months; it also lacks real time distribution system kinetics that are essential to predicting the temperature-driven episodes of elevated DBP levels (Malcolm Pirnie Inc., 1999).

The potential for potentially significant underprediction of the number of locations with Stage 2 compliance needs will lead to EPA's underestimation of the costs and overall financial impacts of the rule. It will also contribute to a potential understatement of risk reductions and associated benefits (all things equal). However, the occurrence issues alter both baseline and post-rule estimates, so the overall net change in potential risks and benefits due to the rule as proposed is not certain.

4. Exposure Assessment

Because EPA's benefits analysis is driven primarily by population attributable risk (PAR) estimates derived from epidemiological research (see Chapter 5), there is limited explicit use of drinking water consumption patterns (i.e., liters per day) involved in the Agency's risk assessment and associated benefits analysis. Instead, the PAR-based approach implicitly assumes that the level (or distribution of levels) of tap water consumption in the general population of 2008 and beyond will be the same as in the study populations (and time periods) from which the PAR estimates are formed. This becomes a core issue with respect to if or how the epidemiological evidence for DBPs and cancer can be applied to the population of the future affected by the regulation.

In addition, tap water consumption issues have other important implications for some aspects of the benefits analysis, and these also are addressed in this chapter. For example, the toxicology-based risk estimates are based on multiplying percentage changes in estimated finished water DBP concentration levels (in micrograms per liter) by an assumed number of liters of distributed CWS water a person is expected to drink in a day. Also relevant to this review is the issue of the duration of exposure, since the carcinogenic risks that may be associated with DBPs are expected to arise from long periods (e.g., 30 years or more) of exposure, and (based on epidemiological evidence presented by EPA) may be positively correlated with duration of exposure.

4.1 Daily Water Consumption: Data from USDA

In the limited use of toxicological evidence applied in the EPA Stage 2 EA, baseline levels of bladder cancer cases, and reductions in these cancers, are based on an assumed level of water consumption of 2 L/day, over a 70-year (i.e., presumed lifelong) duration of exposure (U.S. EPA, 2003a, Exhibit 5.8). These are standard risk assessment assumptions. However, these are unrealistic levels and durations of exposure for virtually all Americans. Therefore, when these conventional precautionary assumptions of exposure are used in a benefit-cost context, they overestimate lifetime exposures and, hence, likewise overestimate the level of risk. Given EPA's use of linear dose-response functions, both the baseline risks and the projected risk reductions attributed to the proposed Stage 2 options are overstated by the same proportion that lifetime exposures are overstated.

In terms of daily tap water intake, EPA typically has used data from the U.S. Department of Agriculture's (USDA's) *1994 — 1996 Continuing Survey of Food Intakes by Individuals* (CSFII). These data have been used for the proposed LT2 rule, for example (U.S. EPA, 2003c). The USDA distribution data based on water ingestion from CWS sources have a median (50th

percentile) ingestion level of 0.71 L/day and a 90th percentile of 2.02 L/day. The mean from this distribution is 0.93 L/day. Therefore, the average level of CWS water ingestion (0.93 L/day) is less than half the level used in EPA's toxicological risk analysis (2 L/day), and half of the population consume roughly one-third or less than this amount (<0.71 L/day, compared to 2 L/day). Accordingly, EPA's results are likely to overstate the number of cases avoided by a factor of more than 2, based on daily CWS tap water intake alone (all else equal).

Second, there are several reasons to believe that even the results from the USDA distribution of CWS water use may overstate daily ingestion levels for some DBP-relevant waters. The USDA water use estimates include both *direct* and *indirect* water consumption. Indirect water is used for final food preparation in the home, and includes water used for coffee and tea, beverages for which the water is boiled or steamed before ingestion (U.S. EPA, 2000a). Hot coffee and tea preparation therefore should volatilize TTHMs (but not HAA5), thereby removing them from the ingested water. Therefore, a portion of the USDA estimates of ingested waters would not carry TTHM-related ingestion risk (although the volatilized TTHMs might pose some inhalation-related risks). Mean per capita direct use of CWS water is 0.50 L/day (U.S. EPA, 2000a), which is 54% of the total water use mean of 0.93 L/day (and only 25% of the 2 L/day EPA uses in its analysis). Thus THM-related risk estimates based on total water consumption could be overestimating actual ingestion risk.[1]

Third, when these USDA data were collected, approximately 13% of daily water intake was from bottled water (U.S. EPA, 2000a). In the 7 to 9 years since the USDA data were collected, there has been a considerable and well documented growth in the use of bottled water and in-home filtration devices [i.e., use of tap water alternatives has grown at a 10% annual rate, according to the Water Quality Association (WQA, 1999)]. Therefore, tap water ingestion for many households has declined considerably since the USDA data were collected. A recent EPA-sponsored Gallup poll (U.S. EPA, 2003c) and AwwaRF-sponsored research by Raucher et al. (forthcoming) indicate that as of 2002, 75% of Americans drink bottled water, 14% to 20% drink ONLY bottled water, 37% use in-home filtration devices, and only 49% to 56% drink exclusively tap water in their homes.

Finally, the ingestion levels for the USDA distributions as presented by EPA in the LT2 EA (U.S. EPA, 2003d, p. 5-23) are higher than applied or derived previously. For example, EPA's prior estimate of tap water intake for the arsenic rule (U.S. EPA, 2000b) reflects a mean of 1.0 L/day, and the National Research Council assumed 0.6 L/day ingestion for its radon risk analysis (NRC, 1999). Tap water intake estimated and reported by Roseberry and Burmaster (1992) is 0.73 L/day at the mean, which also is considerably lower than the levels that EPA now

1. Ingestion risks for HAAs and other unspecified nonvolatile DBPs in the mixture would not be affected by this distinction between direct and indirect ingestion.

reports and applies. This may in part reflect differences in accounting for direct versus total (direct plus indirect) use of tap water.

Based on the above, it seems very likely that the toxicologically based risk and benefit estimates derived by EPA are overstated considerably because of the Agency's assumption of a 2 L/day ingestion level. Direct intake of CWS water is more typically 25% of the EPA-applied level (0.5/2), and total (direct and indirect) use is probably 47% or less of the standard assumption applied by EPA (0.93/2). These percentages would be further reduced by factoring in the considerably increased use of bottled water and in-home treatment devices since the mid-1990s, when the USDA data were collected.

4.2 Duration of Exposure

Cancer risks are generally associated with relatively long, chronic exposures to a cancer-causing compound. The epidemiological evidence presented in the EA suggests this is the case for DBPs in drinking water. To the extent associations are found in these studies between chlorinated surface water use and cancer, they are evident for exposure durations of 30 or 60 years or more (depending on the study), and risk levels appear to increase with the duration of exposure according to many of the cited studies (e.g., see summaries in Exhibit 5.6 in U.S. EPA, 2003a).

Because most of EPA's primary risk estimates are PAR-based, the EA shows no explicit duration of exposures used in the Agency's primary set of benefits estimates. However, there are some implications for how to interpret the PAR-based findings. If one assumes a given PAR estimate is reliable and valid for the study population, then applying that PAR estimate to the general (nonstudy) population will yield unbiased results insofar as the demographics and patterns of exposure across relevant population groups (e.g., by gender and smoking status) are the same in the study group as in the general population. There may be no reason to expect differences between the study and general population in terms of gender, but there may well be important differences in other risk-relevant factors such as duration and levels of exposure, and these are discussed below. (In addition, because tobacco use is declining relative to levels found during the study period — which goes back several decades — smoking status may be another factor that introduces a bias between the study group from the past and the general population of today and the future.)

In terms of duration of exposure, the study groups were specifically selected (at least in some instances) for their relative stability in terms of residential location. Epidemiological study populations were targeted in part because they provided long residential tenancies (reduced confounding) and above average opportunities to evaluate relatively long exposure durations (e.g., Cantor's use of Iowa-based study populations; see Raucher et al., 2001). In contrast, the general U.S. public has been more mobile, and probably increasingly so as we look forward to

the implementation period for the proposed Stage 2 rule. For example, the median duration of residence in the United States is only 5.2 years, and only 2.5% of the American population remains in the same residence for 40 or more years (Hansen, 1998).[2] Therefore, it is plausible and indeed likely that the durations of exposure in several, if not all, of the key epidemiological studies (both in terms of the study location and the past decades that constitute the time period of the study) over-represent high duration exposure periods relative to the general U.S. population (i.e., nationwide, and in the future) as relevant for the proposed rule.

In addition to exposure durations, it is likely that levels of daily tap water ingestion have decreased (and will continue to decrease) over time compared to tap water ingestion levels during the long-past decades of the epidemiological study periods. Increased reliance on bottled water and in-home treatment have already reduced the percentage of households facing elevated CWS-related DBP exposures, and this trend is projected to continue. For example, the 14% to 20% of U.S. households that no longer drink any tap water should be removed from the risk assessment and benefits analysis entirely. This would reduce the national risk and benefits estimates by the same amount (given the near-linearity of the risk and benefits functions applied). In other words, this indicates national results offered by EPA are overstated by 16% to 25% (i.e., 1.0 - 1.0/0.8 = 25%) just because of exclusive bottled water drinkers alone. And, because use of bottled water is growing rapidly in the United States, even higher percentage adjustments may be suitable to reflect the compliance period of 2008 and beyond.

In the context of EPA's limited use of the toxicology-based risk data, the duration of exposure becomes an issue as well. EPA's toxicology-based risk and benefit estimates are based on 70 years of exposure (U.S. EPA, 2003a, Exhibit 5.8). As noted above, residential tenancy is likely to be for periods appreciably less than 70 years, and in fact only 0.6% of Americans are believed to reside in the same location for 70 years or more (Hansen, 1998). While relocated households may face DBP exposures in many of their ultimate residential locations, the levels and composition of the DBP mixture are likely to change considerably over time and location. Thus, the toxicology-based results developed by EPA based on 70 years of exposure are likely to overstate actual exposures and, hence, risks (and benefits) will be overstated as well.

2. When a household exposed to relatively high DBP levels in one location moves to another home, they may or may not face relatively elevated DBP levels again. However, the levels and composition of the DBP mixture are likely to vary across household locations, even if they move within the same CWS. Thus, the nature and level of DBP exposures are likely to change — perhaps considerably — with most residential moves. Because the epidemiological studies rely on relatively stationary populations and reveal cancer risks that tend to be strongly associated with duration of exposure, the changes in exposure levels or mixtures expected in more typical U.S. residents may render the epidemiological results of questionable applicability to the proposed Stage 2 rule.

4.3 Exposure Issues for Sensitive Subpopulations

Two population groups may be at heightened risk from elevated DBP levels in distributed CWS waters. Older male smokers appear most susceptible to bladder cancer, and there may be no specific practical approach to limiting exposures across this particular group (although declines in smoking rates may be highly beneficial).

The second sensitive subpopulation may be pregnant women and their developing fetuses. As noted in the next chapter and in other reviews, the scientific evidence associating elevated DBP levels to possible reproductive and developmental effects is still inconclusive and preliminary. However, if there is indeed an association between some DBPs and fetal loss (spontaneous abortions and stillbirth) or other reproductive/developmental effects, then women of child-bearing age who may be pregnant may wish to take extra precautions.

Pregnant women are encouraged to consume close to 2 L water per day [eight 8-ounce glasses are recommended per day by the March of Dimes (2003), amounting to 2 quarts or 1.9 L]. This is roughly twice the mean level across the population. Pregnant women, who as a class typically receive medical attention and advice, may be warned to avert possible exposures to DBPs in tap water, or might opt to use tap water alternatives regardless of any advisories. If this were true, then exposure levels amongst this group may be reduced in part by averting behavior. However, at this time we have no data to assess this possibility, and a search of various medical and pregnancy-related web sites does not reveal any generic warnings or advisories to pregnant women to avoid tap water.

4.4 Relative Source Contributions and Other Considerations

As noted above, volatile TTHMs may pose health risks through inhalation exposure pathways, and some DBPs may pose risks through dermal contact. Thus, the focus on ingestion alone (especially for TTHMs) may understate the total exposure and risk associated with DBP levels in distributed waters. However, to the extent that inhalation and dermal exposure pathways pose sizable risks relative to ingestion, then it also may be the case that these compounds are posing appreciable risks though swimming pools and other out-of-home exposure scenarios that are unrelated to levels of DBPs in distributed CWS waters. If it is the case that ingested CWS tap water has a comparatively low relative source contribution (RSC) for DBPs, then this has implications for how the epidemiological evidence is evaluated.

More specifically, if there is reason to believe that tap water has a relatively low RSC for DBP-related exposure and risks, then this has implications for the reliability and transferability of the PAR estimates from the epidemiological studies. The key issues are (1) whether the epidemiological studies controlled for these other significant exposure pathways

(i.e., confounders), and (2) whether the study populations (in time and location) have similar exposure patterns for these other DBP exposure pathways when compared to the general, rule-relevant population (nationwide, in the future). If either or both of these aspects are problematic (and we suspect that both aspects may be present and pose notable limitations), then the inherent reliability and applicability of the epidemiological evidence to the Stage 2 rule become increasingly suspect. This issue requires additional investigation by EPA and stakeholders.

Finally, exposures may be seasonal in terms of both water ingestion (probably highest in the hotter summer months) and the general seasonal patterns of elevated DBP levels. DBP levels are positively related to temperature, so there is a possible coincidence of highest tap water ingestion rates during periods when DBPs are at relatively high levels. However, DBP formation also is strongly correlated to the levels of precursors (i.e., TOC), which in some locations may peak with spring runoff and in other locations may be highest in fall months because of leaf litter organic materials. Thus, it is possible that the exposures may be especially high if comparatively elevated seasonal DBP levels coincide with high ingestion intervals. For cancer risks, this is not an issue (since the relevant exposure is based on long-term averages), but for potential reproductive risks that may be associated with limited but time-sensitive exposure periods, this may be a factor to consider in future investigations.

4.5 Conclusions

A fundamental component of the risk assessment and benefits analysis is the amount of tap water ingested by the public and the duration of those exposures. The discussion above reveals several concerns with the assumptions employed implicitly and explicitly by EPA in this aspect of its EA. For example:

- The exposure scenarios (levels, mixtures, and durations) in the epidemiological studies of past study populations may be far different from what will apply in the general U.S. population affected in the future by the proposed Stage 2 rule. This may significantly reduce the validity of applying the study-based PAR estimates to the Stage 2 rule population.

- Exclusive use of bottled water would reduce the subsequent risk analysis and benefits results by approximately 20%.

- Adjusting exposures to reflect direct ingestion in the home might reduce the THM-related ingestion risk and benefits results by as much as 50%.

- Reflecting possible averting behavior by pregnant women and other sensitive populations would directly affect proportionally an important component of the Agency's concerns over potential reproductive and developmental risks.

While the drinking water intake and duration values used in an analysis are not often the focal point of much scrutiny, the values used do have a sizable impact on the ultimate risk and benefit findings. EPA should reconsider how it has approached this aspect of the EA, and provide improved and more expansive analysis and documentation.

5. Quantification of Health Risk Reductions

EPA develops quantified estimates of the reduction in two types of adverse health effects in the EA for the proposed Stage 2 rule. These health endpoints are bladder cancer (both fatal and nonfatal) and reproductive risks (fetal loss, through either spontaneous abortion or stillbirth).

For both types of adverse health endpoints, the Agency relies on epidemiological data to develop estimates of population attributable risk (PAR). These PAR values are then used to develop empirical estimates of the projected number of cases at baseline (i.e., without the preferred option for Stage 2). This is accomplished by multiplying the PAR estimates times the total national number of annual cases of each relevant adverse health outcome. The Agency then estimates the change (reduction) in case numbers that EPA anticipates would be attributable to implementation of the preferred Stage 2 option, based on projections of exposure (dose) reductions.

For both types of health endpoints, there are considerable concerns with the manner in which EPA has developed its quantified risk estimates from the underlying epidemiological evidence. In this chapter, an overview is provided of the core issues and concerns (additional and more detailed critiques of these issues are being developed by other reviewers).

5.1 Bladder Cancer

A large body of epidemiological research conducted over the past several decades has explored potential statistical associations between the duration of past exposure to drinking waters drawn from chlorinated surface waters, and the incidence of various cancers (especially cancer of the bladder). The results from this body of research tend to suggest that such an association might exist.

5.1.1 A suitable association between bladder cancer and DBPs is lacking

The critical question that arises is whether the findings from the existing body of epidemiological studies might logically and defensibly infer an association between DBPs (notably THM4, HAA5, or other DBPs likely to be reduced through compliance efforts) and bladder cancer. This is where the EPA analysis makes a fundamental and unsubstantiated leap of faith that undermines the credibility of its approach and findings. Quite simply, the epidemiological studies of *duration of chlorinated surface water use* do not support an association between the proposed Stage 2 regulatory target (*levels of THM4 and HAA5*) and a reduction in bladder cancer incidence.

EPA came to this conclusion itself, when developing the Stage 1 DBP rule, in 1998. At that time, given the inability to properly interpret the available epidemiological evidence as providing a basis of association, EPA developed empirical estimates simply as an illustration of the potential magnitude of benefits (using an illustrative range of PAR estimates of 2% to 17%).

For the Stage 2 rulemaking, there is no fundamentally new evidence upon which to assess the bladder cancer risks (aside from a meta-analysis that shares the limitations of the original, underlying studies). Nonetheless, EPA departed from its prior position of considering the evidence suitable only for illustrative sensitivity analysis, and the Agency is now using the same evidence to formally quantify regulatory benefits. This is inappropriate. The available epidemiological evidence is still inadequate for the purposes for which it is applied by EPA in the Stage 2 EA.

5.1.2 The linear dose-response function is fundamentally inconsistent with the cessation lag analysis

In its application of the PAR values it derives from the epidemiological evidence, EPA implicitly applies a linear dose-response function for DBPs. The toxicological evidence interpreted by EPA also typically assigns linear dose-response functions for DBPs (chloroform being a notable exception). The linear dose-response functions are compatible with a carcinogen that acts as an initiator. In contrast, a carcinogenic compound that is a promoting agent will have a highly nonlinear dose-response function.

For its cessation lag analysis (see Chapter 6 for further discussion), EPA uses estimates based on tobacco smoke and lung cancer. The carcinogenic agents in cigarette smoke are dominated by promoters, especially with respect to cessation lag (Crawford-Brown, 2003).

Therefore, EPA has developed and applied an internally inconsistent and biologically implausible set of assumptions about DBPs. DBPs must either be treated as initiators (linear dose-response functions and relatively long cessation lags) or be viewed as being dominated by promoters (and have highly nonlinear dose-response functions and short cessation lags). EPA cannot have it both ways. Accordingly, EPA should clarify its view on the mode of action for DBPs and apply that perspective in an internally consistent, biologically plausible manner.

5.1.3 The epidemiological evidence is inconsistent with the toxicological data

There is a considerable difference between the number of excess cancer cases anticipated because of DBPs when one uses EPA's interpretation of the epidemiological data and results derived when applying unit risk factors derived from the toxicological evidence. The EPA-estimated reductions in bladder cancer cases due to the proposed Stage 2 rule are:

- between 20.1 and 182.2 cases per year when the PAR-based analysis is applied
- between 1.7 to 4.0 cases avoided when the toxicology-based results are applied.

The results thus differ by a factor of 11.2 to nearly 45.6 (e.g., 182.2/4.0 = 45.6). A key issue that thus arises for the benefits analysis is how to compare the epidemiological results (for bladder cancer) against what is known about the risks of individual DBPs from the body of toxicological research.

Crawford-Brown (2003) discusses these inconsistencies and points out an additional way to compare the epidemiological and toxicological bases for risk estimates by drawing on the results of the epidemiological study by King and Marrett (1996). That study indicates a slope factor of 0.11 per 1,000 µg/L TTHM-year, and if one assumes 73 years of exposure, this translates into a lifetime risk factor of 8E-3 per µg/L TTHM. Comparing these results to the toxicological information requires an assumption about the average ratio of the TTHMs in the water. Even if the mixture is assumed dominated by bromate (with the highest Unit Cancer Risk Factor), the epidemiological value is higher than the toxicological value by a factor of 400 (i.e., 8E-3/2E-5 = 400). Given that bromate is not likely to dominate any DBP mixture (especially where ozonation is not used, as would be the case in the past periods covered by the epidemiological studies), the magnitude of the difference is probably even higher than 400-fold.

Crawford-Brown (2003) further points out that the Agency argues that this difference may arise because so many DBPs are of unknown toxicity, and so the risk calculated on the basis of the summation of risks from individual DBPs can at best be considered an underestimate. EPA also argues that exposure through drinking water may include inhalation and dermal contact, routes not considered in the toxicological studies. Still, the epidemiological results seem to be significantly out of line with what is known from toxicological studies, raising the distinct possibility that the epidemiological studies are compromised by one or more significant confounding factors associated with exposure to chlorinated surface water.

5.1.4 The role of tobacco smoking is fundamental to assessing bladder cancer risk

There is a well-recognized link between tobacco smoking and the incidence of bladder cancer. As the EPA Science Advisory Board has previously noted, “A well established and major risk factor for lung and bladder cancer is smoking. Since this falls into the area of common knowledge, we will not belabor the point further” (U.S. EPA, Science Advisory Board/DWC, 2000a).

In accounting for the bladder cancer risks posed by DBPs, the confounding effects of tobacco smoking need to be addressed and, if possible, accounted for. This adjustment is potentially significant. This can be deduced by noting that in the United States, smokers are 2 to 3 times more likely to develop bladder cancer than nonsmokers (NCI, 2000), and that some 48% of

bladder cancer cases are realized in smokers (NCI, 1998), even though smokers constitute only about 23% of the U.S. population (CDC, 2003).

In the various epidemiological investigations that have associated elevated bladder cancer rates with exposure to drinking waters drawn from chlorinated surface water supplies, the elevation in risk appears to be associated almost entirely with smokers, and particularly male smokers. As noted by Crawford-Brown (2003), "There is a possibility, therefore, that the risk is confined to male smokers, with smoking being the underlying cause of an increased susceptibility to D/DBPs."

While scientific evidence at this time does not clearly show whether there is a synergistic or other mechanism that accounts for the prevalence of bladder cancers among smokers, the association is strongly evident. Given that tobacco smoking is a voluntary activity, it can be argued that the incremental elevation in bladder cancer risk in smokers relative to nonsmokers should not be considered in estimating baseline cases attributed to DBPs (or to the reduction in cases estimated due to the proposed Stage 2 rule). EPA should investigate the extent to which the estimated risk reduction (number of cases avoided) would decrease if a "net of smoking" basis was used in its analysis. This would be consistent with overall public policy of making smokers responsible for the added risk they impose on themselves, instead of attributing those voluntarily borne risk increments (and risk reduction benefits) to DBPs (and DBP-reducing policies).

5.1.5 Conclusions

EPA's interpretation and use of the epidemiological evidence for bladder cancer is highly problematic in the context of the Agency's Stage 2 benefits analysis. The evidence does not provide a sufficient basis for estimating reliable PAR values, and the results of the analysis are implausibly inconsistent with toxicological data (and the latter may themselves be overstated due to precautionary assumptions that often are embedded in toxicology-based unit risk factor estimates). The numeric results derived by EPA from the epidemiological evidence should be viewed simply as an illustrative exercise, not as a sound scientific basis for any defensible benefit-cost analysis.

5.2 Fetal Loss

5.2.1 The epidemiological evidence does not support quantified analysis

Epidemiological research on the possible association of one or more DBPs with adverse reproductive effects such as fetal loss is in its formative stages. Some positive associations of adverse outcomes with some brominated TTHMs have been suggested by a small number of

studies (e.g., Waller et al., 1998), leading to a rapidly growing set of other research efforts to examine these issues in greater scope. The body of epidemiological knowledge in this area is still in its infancy, and is mostly useful at this stage for suggesting hypotheses that can or should be tested by more refined and better focused epidemiological studies and toxicological investigations. However, the research at this time is too limited, inconclusive, and inconsistent for drawing any strong inferences about causality or association between DBPs in tap water and reproductive risks.

The FACA Advisory Committee reviewed the research on the potential association between chlorinated drinking water and reproductive effects and, after 2 years of deliberation and consultations with the world's leading experts on the topic, concluded that the evidence was not strong enough to support any empirical estimation of reproductive risks. Because of a lack of reasonably conclusive evidence on causality, the Committee's clear directive to EPA was to NOT quantify these risks (i.e., that quantitative risk assessment was inappropriate given the limitations of the existing body of science).

Despite this clear, explicit, and well reasoned directive from the FACA Advisory Committee, EPA opted to provide an "illustrative calculation" (EA, p. 5-4) of fetal loss reductions (i.e., the number of cases avoided per year) that EPA believes can be attributed to the proposed Stage 2 rule. While EPA may be correct in characterizing the potential association of DBPs with reproductive risks to be a matter of "concern," the Agency's choice to develop and publish empirical estimates of these effects is clearly inappropriate. This exercise by EPA might be construed as evidence that the Agency opted to place political expediency ahead of good science and sound judgment.[1]

1. One example is EPA's use of the estimated odds ratio results from the Dodds et al. (1999) study, as published by King et al. (2000), to derive a PAR value from which to estimate stillbirths. The odds ratio from Dodds et al. for stillbirths was 1.59 for TTHM levels above 80 µg/L versus below 80 µg/L (as reported by EPA, in EA Exhibit 5.2). Only one other study, Bove et al. (1995), was listed as examining stillbirths, and they derived an odds ratio of 0.65 (and a 95% confidence interval of 0.45 to 0.95). Why did EPA opt to use the King et al. (i.e., Dodds et al.) results to quantify stillbirth risks in lieu of also factoring in the Bove et al. findings? While the latter study did not disaggregate THM components, is this a sufficient (or even relevant) basis for disregarding this study when examining THMs in aggregate, as EPA does in its analysis? Whether or not EPA was "cherry picking" only those results that supported high fetal loss estimates, the presence of highly conflicting empirical results reveals the problem of trying to develop quantified estimates from the body of research at this time.

5.2.2 The EPA review of epidemiological evidence is neither complete nor even-handed

A central characteristic of the epidemiological studies cited by the Agency in the EA is that the many epidemiological studies available typically examined multiple endpoints, and dozens of studies were considered. When a large number of endpoints are examined, across multiple studies, it is important to avoid data dredging.

Data dredging refers to the process in which multiple associations are considered, and the ones showing a positive association is highlighted. This appears to the how the Agency reviews and portrays the results of the epidemiological investigations for reproductive effects within the EA. As noted by Crawford-Brown (2003), the problem with such an approach is that random positive associations will appear even where there is no causal connection, so long as a sufficient number of associations are examined and the statistical significance of each association is not given an important role. In the case of the EA, the Agency seems to have simply provided a laundry list of positive associations noted amongst many endpoints and studies, with no systematic exploration of their strength, their specificity, their consistency, or the expected rates of false negatives and positives.

This is not in keeping with EPA's own *Carcinogenicity Guidelines*, nor is it consistent with the much higher standards of assessment typically brought by the Agency to the analysis of epidemiological data. Until a far more balanced and systematic assessment is made of the existing and at times conflicting body of research, it is inappropriate for the Agency to suggest that the complete body of data support the claim that exposure to TTHMs and/or HAAs, at levels expected in drinking water, is associated with increased incidence of the reproductive effects mentioned in the EA. This was in fact the conclusion reached by the FACA panel after two years of deliberations and presentations by international experts in the field.

5.2.3 EPA overlooks other causative agents and pathways

EPA's interpretation of the epidemiological evidence overlooks several important considerations. For example, if DBPs are indeed causal agents for fetal loss, then exposures to pregnant women through exposure pathways other than in-home tap water must be considered. Swimming pools and associated environs may be a source of considerable exposures to disinfectants and DBPs, yet these do not appear to be factored into the empirical illustration developed by EPA.

Second, EPA's analysis (and the underlying epidemiological studies) do not account for what may be significant causal factors for fetal loss and other reproductive or developmental effects. The research field for these adverse health outcomes is still in its early stages of development, so we cannot know what important causative agents may be overlooked and, thereby, lead to an over-attribution (or erroneous association) of cases to DBP levels.

For example, recently published epidemiological research (Nielsen et al., 2001; Li et al., 2003) reveals a very strong empirical association between miscarriage and prenatal use of aspirin or nonsteroidal anti-inflammatory (NSAID) drugs (e.g., the widely used pain reliever, ibuprofen). Li et al. (2003) found an 80% increased risk of miscarriage (an adjusted hazard ratio of 1.8, with a 95% confidence interval of 1.0 to 3.2) for NSAID use near the time of conception or during pregnancy. The association is higher for use near the time of conception or for a duration of NSAID use of a week or more. Similar results were obtained for aspirin use. The empirical results generally replicate what Nielsen et al. (2001) observed in Danish women, and suggest a far stronger association between NSAID or aspirin use and fetal loss than the evidence found to date on DBPs and fetal loss (in which, of the four relevant DBP studies, the odds ratios are 0.65, 1.06, 1.29, and 1.59, and the 95% confidence intervals include three lower bounds less than 1.0 and upper bounds ranging from 0.95 to 2.1). The link between fetal loss and NSAID and aspirin use may be an important confounder (omitted significant variable) in the inconclusive body of evidence on DBPs and fetal loss.

Other causative factors associated with fetal loss also are not fully or properly considered. These include the potential impact of maternal age (as associated with delayed childbearing for many women in today's society), and chromosomal abnormalities (which may account for more than half of spontaneous abortions, Bick et al., 1999). In addition, there are some key definitional issues that further call into question the reliability and validity of the EPA approach, including how spontaneous abortion has been lumped together with stillbirth in EPA's notion of "fetal loss" even though they are distinct phenomena, and in how fetal loss is reported in the underlying aggregate data from Ventura et al. (2000) versus how it is defined in the epidemiological studies EPA applies to those data (e.g., Waller et al., 2001 use the twentieth week of pregnancy as a cutoff, whereas the Ventura et al. data reflect full term outcomes).

Finally, a basic review of national statistics is revealing. With the promulgation of the original TTHM rule in 1979, EPA initiated the first regulatory controls that limited DBP exposures. If a strong association existed between DBPs and adverse pregnancy outcomes, perhaps some evidence would be revealed in the years since compliance with the TTHM rule was required. Between 1979 and the mid-1980s, when systems were coming into compliance, the national rate of fetal loss ranged from 13.9 to 14.1 per 100,000 pregnancies (Ventura et al., 2000). Since 1988, that rate has climbed to levels ranging from 15.4 to 17.2 per 100,000 (Ventura et al., 2000). Thus, the rates of fetal loss have increased fairly appreciably since the initial efforts to monitor and control TTHM levels. This casual observation is not intended as empirical evidence that fetal loss is not associated with DBP levels; however, it is suggestive that other factors may be far more important as potential causal agents (e.g., advancing maternal age, caffeine use, or aspirin use). The omission of these other causal agents from the epidemiological research on DBPs (and the empirical application by EPA thereof) makes the Agency's quantitative use of the existing DBP evidence fairly weak, if not downright ill-advised.

5.2.4 EPA's interpretation of the ICR data to quantify fetal loss avoided by Stage 2 is flawed

EPA's effort to quantify the potential fetal loss risk reduction benefits is driven by the Agency's estimates of how many location-specific levels of TTHMs above a key threshold are eliminated by Stage 2. The Agency focuses on location-specific quarterly THM4 levels that exceed 80 µg/L as its benchmark, based on its review of selected epidemiological studies (as described in EA Appendix G).

EPA's assessment of locational observations above its THM4 threshold of 80 µg/L is based entirely on the Agency's limited evaluation of the ICR data. EPA's use of ICR data to quantify fetal losses avoided is neither transparent nor reproducible. An appendix to this critique provides specific details.

EPA identifies 39 plants from the ICR database that it believes are already in compliance with Stage 1 (RAA < 80 µg/L for THM4, and RAA < 60 µg/L for HAA5), but are not in compliance with Stage 2 as proposed. This is shown in EA Exhibit 5-10 (262 plants minus 223 plants = 39). These 39 plants thus are the focal point for EPA's Stage 2 analysis (representing plants in compliance with Stage 1, but still not, in EPA's estimation, in compliance with Stage 2).

EPA's analysis is then driven by the assumption that after Stage 2, these 39 plants will have the same proportion of locational quarterly observations of THM4 above the 80 µg/L benchmark as found in the 223 ICR plants that EPA believes to be already in compliance with Stage 2. According to EPA's interpretation of the ICR data (Exhibit 5-13 in the EA), in Stage 1 compliant systems, 7.1% of the location-specific quarterly observations exceed the 80 µg/L threshold, whereas only 2.2% of the locational quarterly observations exceed that benchmark in the ICR plants that EPA considers already in compliance with Stage 2.

Based on this comparison of different plants in the ICR database, EPA asserts that there will be 69% fewer quarterly location-specific THM4 observations above the 80 µg/L threshold after Stage 2 than in Stage 1 (7.1% - 2.2% = 4.9%, and 4.9%/7.1% = 69.0%). Thus, EPA calculates that Stage 2 would reduce the number of fetal losses by 69% from the pre-Stage 2 baseline (i.e., post-Stage 1). This leads to EPA's quantified estimate of 1,100 to 4,700 fetal losses avoided annually because of Stage 2 as proposed (i.e., 69% of the EPA estimate of 1,583 to 6,780 cases that the Agency attributes to chlorinated drinking water post-Stage 1). Note that the EPA baseline estimate for post-Stage 1 may be overstated considerably, but that issue is not explored in depth here.

Independent re-evaluation of the ICR data yields very different results

Because EPA's quantified fetal loss estimates are driven by its interpretation of the ICR data, and because EPA's approach to the ICR data seems misguided (i.e., simply assuming all plants post-Stage 1 will have outcomes like those ICR plants the Agency believes are already compliant with Stage 2), we have taken a closer look at the underlying ICR data for the 39 plants that EPA identifies as representing the facilities that would need to undertake compliance activities to move from Stage 1 conformance to compliance with the proposed Stage 2 rule. This plant-specific evaluation of the ICR data is much more relevant to the regulatory analysis than is the EPA approach, and it reveals very different conclusions about the degree to which Stage 2 would reduce location-specific quarterly observations below the 80 µg/L THM4 threshold applied in the EPA fetal loss analysis.

Of the 39 plants, one plant (ICR WTP ID #753) has an apparent data entry error (a THM4 level at site AVG1 is reported as 300, whereas the true value apparently was 30.0). Accordingly, this plant is dropped from the analysis. The highest THM4 level recorded for this plant is 52.0, and the highest level at the other three locations in the quarter in question is 30.3. Thus the data error is apparent, and removing this plant does not alter the outcomes developed below.

Examining plant- and location-specific data for the utilities affected by Stage 2

From the remaining 38 plants relevant for Stage 2 compliance evaluation, we had 151 locations within the distribution systems with THM4 observations by quarter (38 plants times four ICR monitoring locations per plant, less one location lacking observations).[2] The ICR data for these 38 plants also furnished 142 plant-level quarterly observations (38 plants times 4 quarters = 152 plant quarters, less 10 missing plant quarterly observations). From these 142 plant quarters and 151 locations, we had 541 total observations of location-specific quarterly THM4 measurements (142 plant quarters times four locational monitoring sites per plant, less 27 missing location-specific quarterly data points).

From the available data for these 38 ICR plants, we counted how many quarterly location-specific THM4 observations were above the 80 µg/L threshold. We noted 65 location-specific quarterly observations above 80 µg/L, out of 541 observations (65/541 = 12.0%), at the Stage 2 baseline (i.e., post-Stage 1).

2. The missing location was the DSE site in plant ICR WTP ID 164.

For each of the 151 locations with ICR data, we also estimated the LRAA for that site. Only four locations exceeded the proposed THM4 LRAA of 80 µg/L.[3] These four locations had a total of eight quarterly observations above the 80 µg/L threshold. Utility compliance efforts, if aimed at 64 µg/L for THM4 (i.e., including the 80% safety factor) would probably drive six of these locational quarterly observations below the 80 µg/L threshold.[4] This would leave 59 location-specific quarterly observations above 80 µg/L (65 at Stage 2 baseline, minus 6 eliminated due to compliance with the proposed Stage 2 LRAA).

After Stage 2 compliance, there would be 10.9% of the quarterly location-specific observations above the threshold (59/541 = 10.9%). Comparing the results between pre- and post-Stage 2 compliance, there is a 9.2% reduction in sites with observations above 80 µg/L (12.0% minus 10.9% = 1.1%, and 1.1%/12.0% = 9.17%).

Alternatively, if one assumes that utilities will also take compliance action at sites where the LRAA is below the proposed MCL of 80 µg/L, but within 90% of the rule (i.e., above 72 µg/L), then four additional sites enter the analysis. Across these sites, there are five additional individual locational quarterly observations above 80 µg/L, of which three are likely to be reduced below 80 µg/L by compliance with 80% of the proposed LRAA-based MCL (i.e., targeting for post-compliance THM4 levels of 64 µg/L). These would leave 56 location-specific quarterly observations above 80 µg/L (59 minus 3 = 56), or 10.4% (56/541 = 10.4%). Comparing the results between pre- and post-Stage 2 compliance under this scenario, there is a 13.3% reduction in sites with observations above 80 µg/L (12.0% minus 10.4% = 1.6%, and 1.6%/12.0% = 13.3%).

Therefore, when examining DBP monitoring results for the 38 specific plants and associated locations in the ICR database that EPA identifies as representing systems moving from Stage 1 to Stage 2 compliance, the anticipated change in site-specific locations with quarterly observations above the 80 µg/L threshold is far lower than EPA claims when it looks at the data at the more aggregated level. Instead of EPA's claim of a 69% reduction in location quarters with THM4 values above the threshold, the ICR data actually reveal the true reduction is likely to be closer to 9% to 13%. Thus, even if one accepts EPA's estimate of DBP-associated fetal loss at the Stage 2 baseline, the resulting quantified estimate of reduced cases is overstated by a factor of 5.2 (69.0/13.3) to 7.5 (69.0/9.2).

---

3. The LRAAs were 80.7 µg/L (MAX site for plant 113), 83.3 µg/L (MAX site for plant 124), and 97.6 µg/L (AVG1 site for plant 266), and 81.9 (MAX site for plant 605). Two of these four sites were missing samples for low THM4 quarters (these appear to be low quarters for THM4 based on other plant location data). Therefore, inclusion of any missing quarters may have led to LRAA < 80 µg/L for plants 113 and 605.

4. Based on estimating the percent reduction required to move the LRAA value from its reported level to 64 µg/L, and applying that percent reduction to each quarterly observation.

Accounting for the composition of THM4 levels and reductions

The above analyses are based on THM4, of which chloroform ($CHCl_3$) is the dominant compound. At the median, chloroform accounts for 70% of the THM4 levels observed in the ICR for the 65 locational quarterly observations above 80 µg/L. However, chloroform appears to be the least likely candidate of the TTHM mix for possible association with reproductive or developmental effects. It is the brominated THM species that are most suspect, given the limited available data, as the DBP agents possibly associated with these adverse outcomes.

Looking more closely at the ICR data for the four plants in the relevant Stage 2-impacted set (i.e., plants with a site-specific LRAA for THM4 above the 80 µg/L proposed MCL), there are a combined 8 site quarters (out of 14 quarterly data points) with a location-specific THM4 levels observed above the 80 µg/L benchmark. These are the "peaks" that drive EPA's quantitative fetal loss illustration.

Next, based on the existing epidemiological evidence provided by Waller et al. (1998), one might focus on BDCM as the specific THM4 species with the greatest likelihood of being the potential causal agent in any elevated risk of fetal loss, especially at levels above 17 µg/L (a level adopted here from Waller et al. as an illustrative benchmark). Among the 4 relevant Stage 2-impacted sites (i.e., sites with THM4 LRAA > 80 µg/L), there are 14 quarters of data, of which 5 have location-specific quarterly observations with BDCM at or above 17 µg/L.

Assuming that the impacted facilities aim for Stage 2 compliance at an LRAA of 64 µg/L for THM4 at these 4 sites, and also assuming that each compound in the THM4 mix is reduced by an equal percent,[5] then the proportional decrease in BDCM would drive only one of the five quarterly observations for BDCM from above 17 µg/L to below that benchmark level (i.e., a post Stage 1 BDCM level of 22.0 becomes 16.9 µg/L post Stage 2). Thus, the baseline of 5 of 14 site quarters in regulation-impacted systems would be altered due to the regulation to 4 site quarters with BDCM above 17 µg/L. This is a reduction of 7.1% (1/14 = 7.1%) in BDCM quarters above

5. It is not known how the compliance-related reductions may be split across the 4 species that constitute THM4. Several factors, including source water characteristics and the control strategy implemented, will determine the comparative reductions across species. However, THM4 reductions may, in many locations, result in a higher proportion reduction in chloroform than in the brominated species (e.g., where controls result in organic content being reduced and/or chlorine dose being reduced, while bromide levels are unchanged).

17 µg/L amongst the system sites that are out of compliance with the proposed 80 µg/L LRAA for THM4.[6]

A more suitable comparison to the ICR data interpretation developed by EPA is to look at the observations for all 38 plants that drive the Stage 2 fetal loss analysis (as described in previous sections). From these 38 plants, we have 539 location- and quarter-specific observations for BDCM levels. Out of the 539 data points for BDCM, there are 114 observations with quarterly site levels of 17 µg/L or greater (114/539 = 21.2%).

As noted above (and in footnote 6), only 1 or 2 of these quarterly site observations is projected to move below 17 µg/L based on compliance with Stage 2 as proposed (assuming proportional reductions as needed to get each impacted site's THM4 LRAA down to 80% of the proposed MCL). This would leave 112 or 113 locational quarters above the 17 µg/L BDCM benchmark used here. This amounts to a reduction of from 0.9% to 1.7% in the number of site quarters above the BDCM benchmark, due to the rule as proposed (i.e., 113/539 = 21.0%; 21.2% - 21.0% = 0.2%; and 0.2%/21.2% = 0.9%).

The above illustration — with a focus on BDCM in lieu of the broader THM4 focus — reveals the impact that chloroform may have on EPA's efforts to quantify fetal loss. When looking at the THM4 aggregate in the preceding section, we attributed a 9.2% to 13.3% reduction in baseline DBP-related fetal loss cases to the proposed Stage 2 rule. However, when focusing on BDCM, the reduction attributed to the rule shrinks to 0.9% to 1.7%. If compliance efforts result in preferential reductions in chloroform relative to the brominated species, this impact would be even more sizable (and the fetal loss risk reductions potentially associated with the rule would be smaller).

If one applies the BDCM-relevant ICR observations to the EPA's estimated baseline estimate of fetal loss, then Stage 2 would be associated with a reduction of only 14 to 115 cases per year (i.e., 0.9% of 1583 = 14). This indicates that even if one were to ill-advisedly use the epidemiological evidence to quantify fetal loss, and even if one accepted EPA's estimated PAR values and the associated Agency estimate of baseline fetal loss levels associated with DBPs in drinking water, then a more suitable interpretation of the Stage 2-relevant portion of the ICR data would in and of itself reduce the estimated Stage 2-attributed fetal loss cases avoided to levels between 1% and 2% of the numbers published by EPA.

---

6. Alternatively, if one assumed that any system with a site observation at 90% or more of the proposed MCL (i.e., an LRAA above 72 µg/L for THM4) would make compliance efforts, we then obtain 10 location-specific quarterly observations (out of 26 location-specific quarterly data points amongst 8 sites) with BDCM above 17 µg/L at the Stage 2 baseline. Stage 2 compliance efforts would be projected to drive 2 of these observations below 17 µg/L for BDCM, resulting in a 7.7% decrease due to the rule (2/26 = 7.7%).

5.2.5 EPA's fetal loss analysis may violate federal information quality guidelines

The fetal loss illustration may well be in violation of EPA and Office of Management and Budget (OMB) information quality guidelines (U.S. EPA, 2002; OMB, 2002). Among the reasons this may be the case:

1. EPA itself states that it "does not believe the available evidence provides an adequate basis for quantifying potential reproductive/developmental risks." Yet despite this assessment, the Agency nonetheless proceeds to develop empirical estimates. EPA characterizes these empirical estimates as "illustrative calculations" yet the Agency highlights their numeric results throughout the rulemaking package and EA.

2. The EPA analysis is not objective either in the context of its presentation or in terms of its substance. EPA's presentation of the epidemiological evidence is not even-handed or unbiased, and its illustrative calculation is not accurate, reliable, or unbiased. EPA cherry-picked what data and studies to use and ignored a more balanced weight of evidence approach that would have more fully captured the mixed and inconsistent results evident in the body of relevant research.

3. EPA does not appear to have conducted any substantive expert peer review of the approach or results, nor did the Agency conduct a pre-dissemination review (as clearly required under the OMB guidelines for information quality).

5.3 Conclusions

EPA quantitative estimates of the number of bladder cancer cases and fetal losses avoided per year due to the Stage 2 rule are the results of ill-advised applications of weak empirical evidence from the body of available epidemiological research. Using these studies as a basis for quantifying benefits is not consistent with good science.

In the case of bladder cancer, there is still no established causal mechanism, which would be needed as a premise for developing and applying PAR estimates in a manner that is scientifically justified. The results are also so appreciably inconsistent with the toxicology-based evidence that it is quite apparent that the estimates are not likely to be credible.

For fetal loss, the epidemiological evidence is even thinner and more inconsistent, and EPA has been rather selective in what portions of the research body it opts to use, and the manner in which they are applied. Further, the quantitative fetal loss estimates are developed by EPA using a very unsuitable interpretation of the ICR data. Even if the epidemiological evidence had been sufficiently robust to justify EPA's efforts to develop quantitative estimates of fetal loss, the

Agency's application of its poorly conceived PAR estimates to the ICR data yields very misleading results.

While the scientific evidence may be such that a "concern" can be noted and some exposure-reducing options considered, EPA's development and application of quantitative estimates of cases avoided is unfortunate and probably misleading. Steps may be warranted to reduce levels of DBPs (or, more specifically, to reduce the high-side variability in DBP levels across locations within a distribution system), even if the associated health benefits are not immediately quantifiable. However, we strongly object to the Agency's efforts to quantify specific outcomes when the science does not support such endeavors, which is especially true for the potential reproductive and developmental benefits of the proposed rule.

6. Characterization and Valuation of Bladder Cancer Risk Reduction

Of the potential adverse health effects that could potentially be avoided with implementation of the Stage 2 DBP rule, the EA monetizes only expected reductions in the incidence of fatal and nonfatal bladder cases.

To value the expected reduction in fatal bladder cancers, EPA uses a distribution of monetary estimates of the value of a statistical life (VSL) that has been incorporated in numerous economic assessments of proposed EPA regulations for various media (e.g., air and water). While theoretical and empirical questions continue to surround the development, use, and interpretation of VSL estimates, the precedent for their use is clearly established and their application in context of the Stage 2 rule raises no new issues or fundamental concerns. In contrast, there are no available studies that directly estimate willingness to pay (WTP) to avoid a nonfatal case of bladder cancer.[1]

To address this data gap, EPA relies on estimates from two valuation approaches for nonfatal bladder cancer cases. First a quantitative relationship between fatal and nonfatal lymphoma case values developed in a contingent valuation study (Magat et al., 1996) is used as a linear scalar to adjust VSL estimates. Second, nonfatal bladder cancer cases are valued using WTP estimates for a case of chronic bronchitis, following a direct recommendation from EPA's Science Advisory Board (SAB).

This chapter discusses the valuation of fatal bladder cases, in Section 6.1, to provide background for the discussion of the development of the nonfatal case estimates in Section 6.2. Next, Section 6.3 briefly discusses using estimates of life years saved as an alternative measure for summarizing the expected mortality impacts of the Stage 2 rule. Section 6.4 addresses the incorporation of cessation lag and discounting issues into the economic analysis of the Stage 2 rule.

1. Economists prefer *ex ante* WTP-based estimates of the value for avoiding an adverse health outcome because they are constrained by disposable income but allow the respondent to consider the relative impact and importance of all aspects relating to the diagnosis and treatment of a condition as well as any pain and/or activity limitations that the condition or its treatments may impose over time.

6.1 Valuation of Fatal Bladder Cancers in the Stage 2 Rule's Economic Assessment

To value expected reductions in fatal bladder cancer cases attributable to implementation of the Stage 2 rule, EPA first selects a value from a distribution of the VSL estimates and then adds to it an estimate of the medical costs for treating a fatal case of bladder cancer (U.S. EPA, 2003a).

The VSL distribution is defined by estimates from 26 studies developed using labor market (i.e., wage-risk) studies and contingent valuation surveys. The resulting VSL distribution has been incorporated in analyses of the economic benefits of proposed EPA regulations expected to provide a reduction in the risk of a fatal health outcome for contaminants in a variety of media (e.g., air and water).[2]

While reductions in cancer risks are not a focus of the studies that contribute to the VSL distribution, the appropriateness of using this VSL distribution to monetize reductions in mortality risks based on anticipated changes in the incidence of fatal cancers has been specifically recognized (e.g., U.S. EPA, Science Advisory Board, 2000b). In addition, recent economic assessments of EPA regulations, such as the one for arsenic concentrations in drinking water, have established a clear precedent for incorporating values drawn from this distribution to monetize expected reductions in the risk of fatal bladder cancer cases.

As a result, there are no new substantive issues regarding the incorporation of results from the VSL distribution to value an expected reduction in fatal bladder cancer cases attributable to the implementation of the Stage 2 rule.[3]

Similarly, there is little question surrounding the addition of expected medical costs for treating a fatal bladder cancer case to any selected VSL estimate. In part, this reflects the relative magnitude of the components contributing to the final fatal case values. The lowest estimates in the pool of results that define the VSL distribution are in the $800,000 to $900,000 range, and the estimate of medical costs is in the range of $90,000 to $100,000. Therefore, adding medical costs to a VSL estimate does little to change the overall magnitude of the result. Finally, adding this value to any selected VSL estimate for valuing a fatal cancer case directly reflects the

2. EPA incorporates the VSL distribution in a Monte-Carlo analysis designed to reflect the variability and uncertainty in key inputs to a benefit assessment. The Monte-Carlo design creates a distribution of monetized benefits for the mortality risk reduction by selecting a value from the VSL distribution over many separate draws.

3. There is still a lively debate regarding composition and subsequent use and interpretation of results from the VSL distribution. However, these issues are not directly related to the use of this distribution in the Stage 2 EA, which is consistent with existing EPA precedent.

incorporation of U.S. EPA SAB comments provided for the recent arsenic MCL rule, which also focused on valuing changes in bladder cancer cases (U.S. EPA, Science Advisory Board, 2000b). Therefore, the adjustment is appropriate for the Stage 2 rule given that the same health outcome is being considered.

6.2 Valuation of Nonfatal Bladder Cancers in the Stage 2 Economic Assessment

In contrast to the VSL distribution that provides a reasonably defensible and EPA-approved source of data for valuing risk reductions for fatal bladder cancer cases, no studies provide WTP estimates for avoiding a nonfatal bladder cancer. To fill this informational void for the economic assessment of the proposed Stage 2 rule, EPA uses a linear scalar to adjust values selected for fatal cases from the previously described VSL distribution.

Specifically, Magat et al. (1996) used respondents' results from a computerized risk-trading survey program in a contingent valuation framework to estimate the value of a nonfatal case of lymphoma (cancer of the lymph system) relative to a fatal lymphoma case and other health outcomes (e.g., automobile death, nerve disease). From the respondents' results, Magat et al. (1996) conclude that "the median respondent found reducing the risk of a nonfatal form of lymphoma to be worth up to 0.583 times the value of an equivalent reduction in the risk of an automobile death" (Magat et al., 1996, p. 1129). Therefore, EPA takes 58.3% of a value drawn from the VSL distribution for a fatal bladder cancer case in the Monte Carlo simulation of the Stage 2 rule's economic assessment to value nonfatal bladder cancers. For a VSL of $6.3 million (year 2000 dollars), the implied WTP to avoid a nonfatal bladder cancer is nearly $3.7 million.

On the surface, incorporating a percentage of a VSL estimate to monetize nonfatal bladder cancers appears straightforward. However, because it is effectively a "benefits transfer," the appropriateness of this approach must be evaluated against EPA's benefit transfer criteria (U.S. EPA, 2000a).

6.2.1 Issues with the use of the Magat et al. (1996) results in a benefits transfer

In its *Guidelines for Preparing Economic Analyses*, EPA noted that benefits transfer valuations are most suitable when the following conditions hold (U.S. EPA, Science Advisory Board, 2000b, p. 87):

1. The good, service, or outcome originally valued is essentially the same as what will be valued by the benefits transfer.

2. The characteristics and conditions surrounding the good originally valued and that to be valued by the benefit transfer are similar, including the extent of any changes that are being valued.

3. The populations affected in each valuation scenario are similar.

In evaluating these conditions for the application of the scaled VSL estimates to nonfatal bladder cancer cases in the Stage 2 EA, the essential question is how similar are a nonfatal case of lymphoma, based on the description provided to the study's respondents, and a typical nonfatal bladder cancer case. Ideally, this comparison would look at information for the two types of cancers on patient characteristics, disease treatment, and side effects (e.g., pain, activity limitations). However, it is in developing this information for the nonfatal lymphoma case valued in Magat et al. (1996) where problems arise.

What we know from Magat et al. (1996) is that respondents were presented with a list of consequences of contracting lymph cancer. A distinction was then made between curable lymphoma cases and terminal lymphoma cases based on the probability of death (10% and 100% fatality probability, respectively) (Magat et al., 1996, p. 1125). From reported rankings of the respondents' highest aversion scores for the consequences of lymphoma (both fatal and nonfatal), we know that the following impacts were presented in the original list: bleeding, infections, depression, loss of energy, swelling, fever, weight loss, and sweating (Magat et al., 1996, p. 1127, Table 4). However, comparing the lymphoma aversion results with those reported for nerve disease suggests that lymphomas were essentially presented as a disease that, if not fatal, is survived without long-term consequence. This conclusion reflects that the nerve disease consequences that received the highest aversion scores included "must restrict exercise," "must quit work," "medications required," or "constant pain," none of which appear in the equivalent list for lymphomas.

Unfortunately, this information still provides an incomplete picture of how lymphomas were characterized to the respondents. Critically, it is not clear from the available information if the consequences of lymphomas were presented merely as a list of potential consequences or whether probabilistic information (e.g., likely consequences or consequences observed in some percentage of patients) was also provided. Additional information on the presentation of the lymphoma cases to the study respondents is apparently available in a Duke University working

paper produced by the authors of Magat et al. (1996); however, extensive efforts to obtain a copy of this paper have been unsuccessful to date.[4]

Despite a lack of information regarding the presentation of lymphomas in Magat et al. (1996), we can still provide a limited comparison of bladder cancer and lymphoma consequences in the United States using readily available data from the National Cancer Institute (NCI, 2003). Exhibit 6.1 presents a limited comparison of characteristics and consequences of lymphomas (combination of Hodgkin's and non-Hodgkin's) and bladder cancer.

Exhibit 6.1. Comparison of features of urinary bladder cancer and lymphoma for individuals of all races

Cancer feature	Urinary bladder cancer	Lymphoma (non-Hodgkin's and Hodgkin's)
All age incidence rate per 100,000 (2000 U.S. standard population) for 1996-2000[a]	20.3	21.8
All age death rate per 100,000 (2000 U.S. standard population) for 1996-2000[a]	4.4	9.1
5-year relative survival rate for 1992-1999[a]	81.8%	60.7%
Number of cases, all ages for 1996-2000[b]	34,052	38,020
Number of deaths, all ages for 1996-2000 [c]	58,702	122,228
Median age of patients at diagnosis, all ages for 1996-2000[d]	72	64
Median age of patients at death, all ages for 1996-2000[e]	77	73
Percent of diagnosed cases in persons 65 years and older	71%	48%
Average years of life lost per person dying of cancer all races, both sexes, 2000[f]	10.9	15.3

a. Table I-4, NCI (2003).
b. Table I-11, NCI (2003).
c. Table I-13, NCI (2003).
d. Table I-12, NCI (2003).
e. Table I-14, NCI (2003).
f. Figure I-19, NCI (2003), no figure directly reported for lymphoma. Use information on life years lost for non-Hodgkin's (14.8) and Hodgkin's (24.6) with distribution of deaths from 1996 to 2000 for non-Hodgkin's (115,376) and Hodgkin's (6,852) to get weighted average value = [14.8*(115,376/122,228)] + [24.6 * (6,852/122,228)].

4. The reference to the working paper is provided in Magat et al. (1996) and is as follows: Magat W.A., K. Viscusi, and J. Huber. 1994. "The Death Risk Lottery Metric for Valuing Health Risks: Applications to Cancer and Nerve Disease." Working Paper No. 93-8, (Duke University) Center for the Study of Business, Regulation and Government Policy. Efforts to obtain this paper have included the following: requests to several economists at EPA, requests to the listed Duke University Center, and directly contacting the two surviving authors (Huber and Viscusi). To date none of these efforts has been successful.

Exhibit 6.1 shows that while bladder cancer and lymphoma are diagnosed with about the same frequency in the general population, lymphoma is generally diagnosed in a younger population and has a more significant mortality threat as measured by actual deaths, survival rates, and expected life years lost from a fatality.

In addition, the age differences in patients diagnosed with these cancers are striking. Bladder cancer is almost exclusively an older person's cancer, with over 71% of the diagnosed cases from 1996-2000 occurring in those age 65 and older. In contrast, only 48% of the combined lymphoma cases diagnosed in the same period were in individuals age 65 and older.

In addition, there is a strong potential for these cancers to have different impacts on their patients, at least in the short run, based on a review of treatment descriptions (the comparison was limited to adults with non-Hodgkin's for the lymphomas). In general, it appears that standard treatment for bladder cancer involves surgery or localized infusion of medication into the bladder. In contrast, for adults with non-Hodgkin's lymphomas radiation, therapy or chemotherapy appears to be the central element in standard treatment (Dana Farber Cancer Institute, 2003).

Ultimately, all of this evidence merely raises questions about the extent to which the characteristics and consequences of a nonfatal lymphoma case, as described to respondents in the original Magat et al. (1996) study, are similar or dissimilar to the current characteristics and consequences of a nonfatal bladder cancer case. It is this similarity (or lack thereof) that ultimately will define the appropriateness of the benefits transfer, but this is an area we cannot comment on further without additional information. If EPA intends to use the Magat et al. (1996) findings to value nonfatal bladder cancer, it must provide full documentation of the exact information, context, and wording used in the value elicitation instrument applied in the study. EPA also must then provide a compelling and complete presentation of the suitability of using the Magat et al. findings in a benefits transfer to value nonfatal bladder cancer risk reductions.

Finally, putting aside issues relating to the similarity in nonfatal cases of bladder cancer and lymphomas, it is worth taking note of an interesting conclusion from the comparison of fatal and nonfatal bladder cancer case values provided in Magat et al. (1996). From all available information, it appears that a case of nonfatal lymphoma, as described to the study subjects, would not result in any significant long-term health issues or activity restrictions. In this view, the nonfatal lymphoma represents an adverse health outcome that may seriously disrupt one's health and quality of life for a year or so, after which life would return to normal. Despite this relatively minor disruption, especially in light of the fact that the average age of the respondents is 31.5 years (Magat et al., 1996), a nonfatal lymphoma case still ends up being valued at over half of a fatal case. While there is no direct empirical basis for questioning this result from the Magat et al. study data, it appears to be a very high valuation when considering the relative magnitude of the consequences as presented. While this result may suggest some scenario

rejection and or dread of cancer in any manifestation on the part of respondents, this is less of an issue in the benefit transfer for the Stage 2 rule as another cancer outcome is being valued.

6.2.2 Alternative valuation using the WTP estimates for chronic bronchitis

As an alternative to using 58.3% of VSL estimates to value nonfatal bladder cancer cases, EPA also values these cases by selecting values from a published WTP distribution for avoiding a case of chronic bronchitis (U.S. EPA, 2003a). This amounts to approximately $600,000 (in year 2000 dollars) per case avoided.

Clearly, there is little overlap between the characteristics and consequences of nonfatal bladder cancer cases and incident cases of chronic bronchitis. However, this alternative valuation approach follows the recommendation of EPA's Science Advisory Board, which considered issues relating to the valuation of nonfatal bladder cancers in the recent Arsenic rule (U.S. EPA, 2003a, see footnote 10 on p. 5-66). Incorporating these values is consistent with established precedent and facilitates comparisons of the benefits across rules.

Even though the published chronic bronchitis WTP estimate from the literature has been reviewed by SAB and applied in prior rulemakings (e.g., the arsenic MCL), applying this estimate to nonfatal bladder cancers requires a more critical assessment by EPA. As noted in a previous section, there are standard professional practices and norms regarding this type of benefits transfer. Although these standard protocols for benefits transfer have been incorporated in the Agency's own *Guidelines for Preparing Economic Analyses*, EPA has not adhered to them in this application.

Specifically, EPA needs to consider whether and how the two health endpoints in question may be similar or different in terms of how they affect individuals and, hence, their possible WTP to reduce risks. For example, nonfatal bladder cancer, once in remission, may have little or no impact on the duration or quality of life enjoyed by the afflicted individual. Chronic bronchitis, on the other hand, tends to be a life-long illness that will cause pain and typically result in activity restrictions throughout a person's life. Further, it may also increase the victim's susceptibility to other illnesses and might ultimately contribute to premature mortality by other causes. Thus, it may well be the case that people would hold a lower WTP to reduce a risk of a transient nonfatal bladder cancer than they would for lifelong suffering from chronic bronchitis. This is a classic benefits transfer issue that EPA needs to explicitly explore and evaluate openly in its EA.

6.3 A Useful Metric that EPA Should Provide is Life Years Saved

One approach to quantifying the expected benefits of the Stage 2 rule not incorporated in the economic analysis is the calculation of expected life years to be saved if the proposed rule were implemented. In its most basic application, this could be completed by multiplying the expected fatal bladder cancer case reduction totals by available estimates of average life years lost for the cancer, such as the 10.9 years per case in Exhibit 6.1, as available from the National Cancer Institute (NCI, 2003).

The usefulness of a life-years-saved (LYS) estimate is that it quantitatively integrates information for an outcome of significance (avoided deaths) with characteristics of those affected (e.g., age at diagnosis and death) in a way not currently possible for the mortality risk valuation estimates (i.e., VSL estimates used by EPA do not reflect the extent of expected life duration lost). The LYS estimate thus becomes another measure that can help regulators in making decisions about preferred alternatives for implementing a rule or for deciding between rulemaking options.

One note of caution when considering this approach is to avoid using LYS estimates in conjunction with available VSL estimates to develop monetary estimates of the value of a life year saved. While as a matter of quantification this can be accomplished relatively easily (e.g., incorporating information on the number of expected life years remaining that underlie the VSL estimates, and then calculating a corresponding value per life year using a suitable discount rate), there is currently no consensus within the economic profession regarding the appropriateness of such calculations and their application in regulatory contexts.

6.4 Cessation Lag for Bladder Cancer

The cessation lag portion of the EA is intended to estimate the time lag between a reduction in DBP exposures (i.e., when compliance costs are incurred to reduce DBP levels in distributed waters) and any resulting reduction in cases of bladder cancer. This is used to develop a time path for discounting potential future cancer risk reductions relative to the nearer-term costs incurred for DBP reductions. This section discusses the concepts of cessation lag and its related notion of latency period, and then reviews the manner in which EPA developed and applied a cessation lag to the bladder cancer risk analysis for the Stage 2 rule.

6.4.1 Cessation lags and latency periods are distinct concepts and durations

The term "cessation lag" was coined by the EPA SAB panel reviewing the arsenic rule (U.S. EPA, 2000a), and is used by EPA in lieu of the previously used term and concept of "latency period." The underlying concepts for latency periods and cessation lag are similar — they are intended to reflect the period of time (e.g., years) between a change in exposure and an associated change in the risk of manifesting a specific cancer(or another health endpoint associated with chronic exposure). However, there are some important distinctions to consider as well.

Latency refers to the period of exposure before the cancer is likely to manifest. For example, tobacco smokers may typically accumulate several decades of exposure before lung or other cancers typically manifest (i.e., many smokers begin in their teens or twenties, but associated lung cancers usually manifest in their sixties or later, indicating a latency period of perhaps 40 years or more). Persons exposed to DBPs in drinking water may likewise need to accumulate decades of exposure before any risk of bladder cancer would be evident (assuming the targeted DBPs are indeed a causal agent).

In contrast, cessation lag refers to the period after an exposure has been eliminated (or, presumably, reduced) and the resulting reduction in the risk of manifesting a cancer. For long-term smokers who quit the habit, the cessation lag is fairly short, because lung cancer and other smoking-related risks seem to decline fairly rapidly after smoking is terminated (e.g., for many former smokers, risk levels may approximate those for nonsmokers within 5 or 10 years). Thus, latency periods and cessation lag periods can be very different in duration, and probably depend on the mode of action for the carcinogen of relevance (see below).

6.4.2 Both latency and cessation lags may be applicable to the Stage 2 rule

In a regulatory analysis context, the applicability of latency periods, cessation lags, or both depends on the nature of the policy under consideration and who is likely to be impacted.

For example, a program designed to help prevent teens from becoming smokers will generate benefits in a future period based on the latency period concept. This is because the program is avoiding or reducing exposures for those who have yet to accumulate much (if any) lifetime exposure, and the cancers avoided because of the prevention program would have occurred several decades in the future.

In contrast, a program to encourage and help lifelong adult smokers quit will probably generate benefits within a relatively short time (compared to the teen smoking prevention alternative). This is because the comparatively short cessation lag concept is applied to those who have accumulated high lifelong exposures, since it is their exposures that would be ended.

For the proposed Stage 2 rule, DBP exposure reductions will be spread across the entire U.S. population regardless of age and accumulated lifetime exposure. For those who have accumulated decades of relatively high exposure (e.g., those past median age who have also spent much of their lifetime drinking CWS tap water with elevated DBP levels and who will remain in water systems affected by the Stage 2 rule), the cessation lag concept and associated duration would apply. But for those who are relatively young (e.g., below median age) or those for whom past exposures have been minimal (e.g., older individuals whose lifetime exposures have been relatively low, but then move to a Stage 2-affected CWS later in the life), the latency concept will apply. Thus, some fraction of the cancer risk reduction associated with the rulemaking would accrue according to a timetable associated with the cessation lag, and the remaining portion of the avoided cases would accrue according to the latency period.

In the initial years of rule implementation, cessation lag may be the relatively dominant impact. However, as the years after initial implementation pass, the latency period would become the predominant or sole relevant factor (i.e., as all those with relatively high accumulated lifetime exposures when the rule was first implemented reach the end of the natural lives, and the remaining impacted population is solely those with limited accumulated lifetime exposure).

6.4.3 EPA's approach to estimating cessation lag

EPA applies a relatively short time lag based on its assessment of how lung cancer risks decline after cessation of tobacco smoking (an exponent value of 0.77 is applied in the relevant cessation equation, and this value is based on tobacco smoking data).

EPA justifies its approach by claiming that tobacco smoke and DBPs may both reflect a mix of cancer promoters and initiators. Hence, EPA asserts that the cessation lags might be similar (i.e., short). EPA acknowledges many differences between the smoking and lung cancer scenario, and the potential bladder cancer risks associated with DBP exposure. However, EPA does not develop alternative estimates (e.g., as a sensitivity analysis) or discuss how biased their results might be.

The approach used by EPA results in a projected cancer case reduction of 694 cases within the first 5 years of the rulemaking, compared to an alternative estimate of between 300 and 400 (perhaps less) using more plausible cessation lag exponents. Over a 20-year period, the present value (at 3%) of bladder cancer risk reduction benefits would be 61.8% of the EPA estimate, using an exponent of half the tobacco smoking value applied by EPA. Thus, the monetized value of benefits may be over-projected by a factor of 1.6 (or more), if the true value for the lag period is half or less than the smoking estimate.

6.4.4 Limitations and concerns with EPA's cessation lag approach

EPA provides a fairly good discussion of several of the inherent limitations and uncertainties associated with applying the tobacco smoking cessation data for lung cancer to the Stage 2 relevant scenario of reduced DBP exposures and bladder cancer. For example, EPA acknowledges differences between inhalation and ingestion, cessation versus reduction in exposure, and direct versus indirect pathways to the target organ.

These are all relevant concerns, but numerous additional problems also exist with the EPA approach of using the cessation data for tobacco smokers. Some of the key issues include:

- *Biokinetics:* The carcinogenic agents in tobacco smoke (and their metabolites) have much more rapid biokinetic clearance in the lungs than do DBPs from the bladder; hence one expects a longer cessation lag for DBPs and bladder cancer.

- *Cell turnover rates:* There is a higher turnover rate of lung cells than bladder cancer cells, implying a relatively longer cessation lag where the bladder is the target organ.

- *Initiators versus promoters:* EPA's use of a linear dose-response function is consistent with a belief that DBPs are dominated by initiators. While cigarettes contain a mixture of promoters and initiators, it is the promoters that dominate. If EPA retains its application of the tobacco-based cessation lag, then it is taking a position that it believes promoters dominate the DBP mixtures. Accordingly, EPA would then be compelled to (a) demonstrate this, or at least offer a compelling argument; and (b) apply a nonlinear dose-response function for DBPs, to be internally consistent.[5]

- *Bioaccumulation assumption:* EPA's approach is premised on a key biological argument, which the Agency does not clearly state as its assumption, that there is neither passive nor active bioaccumulation of the cancer causing agent that may be present in the DBP mixture affected by the Stage 2 rule. If there is bioaccumulation, then a relatively short cessation lag is not likely.

- *Cessation versus reduced exposure*: The smoking estimates are based on data from people who ceased smoking (i.e., eliminating exposure), whereas the Stage 2 rule would simply reduce (but not eliminate) exposures to DBPs. Risk reductions are likely to be less

5. Further, in order for any promoting agents in the DBP mix to imply a relatively short cessation lag, the components need to be present at effective doses. These agents are present at effective doses in cigarette smoke, but there are no data suggesting that a promoting agent (or group of promoting agents) is present at effective doses in the DBP levels present in drinking water (Dr. Richard Bull, personal communication, January 2004).

rapid when exposure is not completely eliminated. In addition, if the dose-response function is nonlinear, then reducing (rather than eliminating) exposures may have limited impacts in terms of reducing risks.

It also should be noted that data on bladder cancer cessation lags do exist, but only in relation to the changes in elevated risks of bladder cancer as borne by smokers. The observed cessation lag in those data thus reflect eliminating exposure to the relevant carcinogenic agents in tobacco smoke, and may not have relevance for reduced exposure to DBPs. Hence, it seems inappropriate to draw inferences from the data on smoking cessation and bladder cancer.

6.4.5 Conclusion and suggested alternatives regarding cessation lag period

The cessation lag can have a sizable impact on the benefit-cost outcomes. For example, EPA's present value benefits estimates would be overstated by a factor of 1.6 (or more) for bladder cancer cases if the exponent value is one-half the level EPA applies in the EA.

Given considerable scientific uncertainty regarding the cessation lag, EPA should have, at a minimum, conducted a sensitivity analysis reflecting a range of values (scenarios) that are more plausible than the single value used in its analysis. Lower cessation exponents are more likely than the value applied by EPA.

In addition, cessation lag may be applicable for only a portion of the exposed population, and latency periods would apply to the balance (in lieu of cessation lag). EPA should further explore the use of a blend of both latency and cessation lag periods.

Further, greater scientific scrutiny and peer review are necessary. At a minimum, EPA should pursue independent expert peer review for these issues before applying any smoking-related cessation data for DBPs or other contaminant exposures via drinking water.

Finally, EPA's internal inconsistency between the promoter-oriented assumption necessary to support the short cessation lag estimate and the Agency's implicit use of a linear dose-response function is very troubling. EPA needs to acknowledge this internal inconsistency, and then articulate a logical and defensible position that is both internally consistent and consistent with good science.

6.5 Conclusions

There are several key aspects of how EPA monetizes the benefits of reduced bladder cancer risks in the EA for the proposed Stage 2 rule. These include:

- The use of what appears to be a very high WTP estimate for nonfatal bladder cancer cases, where the benefits transfer is made without following standard protocols to ensure consistency across the study and policy scenarios (i.e., the health endpoints), and where critical documentation needed to make an objective evaluation of the underlying Magat et al. data (e.g., background documentation on the context and survey questions actually posed in the key study) are not available (at least to date) from EPA for public review.

- The use of cessation lag periods that seem to be too short because they are based inappropriately on data for smoking cessation, rely on several questionable implicit and explicit assumptions, are internally inconsistent with the dose-response relationship applied by the Agency, lack proper scientific peer review, may need to be blended with latency periods for portions of the population impacted, and for which alternative (and equally or more plausible) scenarios need to be assessed through sensitivity analyses.

7. Compliance Costs and Net Benefits

This chapter provides a cursory review of EPA's approach and results for estimating the costs of compliance with the proposed Stage 2 rule. The chapter then proceeds to review how EPA compares estimates benefits to costs, with a particular focus on the extent to which an appropriate incremental net benefits perspective is followed by the Agency.

7.1 Compliance Costs

National level compliance cost estimates are not as precise as EPA implies

EPA estimates that the annualized cost of compliance for the preferred regulatory option (80 µg/L LRAA and 60 µg/L LRAA for THM4 and HAA5, respectively, and 10 µg/L for bromate) will be from $59.1 million to $64.6 million (using 3% and 7% interest rates to annualize capital costs, respectively) (U.S. EPA, 2003a, Exhibit ES.4). EPA is fairly certain that compliance costs will approximate these estimates, as reflected in its 90% confidence ranges that are only plus and minus 8.1% of the expected values reported above (e.g., $54.3 million to $63.9 million is the stated 90% confidence range about the "expected value" of $59.1 million, for the 3% interest rate scenario).

We did not conduct an in-depth review of the cost models and assumptions used to develop these estimates, but we find it difficult to believe that EPA is 90% confident that it has accurately forecast the Stage 2 compliance costs within 8.1%. Many complex and uncertain factors underlie the cost analysis, including:

- the occurrence of DBPs pre Stage 1, based on the ICR data
- baseline noncompliance projections, using the SWAT model to project post-Stage 1 DBP levels in distributed waters
- the inherent limitations of SWAT model projections for estimating DBP levels in distributed waters for various compliance options under Stage 2 (especially for estimating DBP levels at the high end of the distribution rather than near the mean)
- the extent to which utilities will apply safety margins (e.g., whether some will aim above or below 80% of the 80/60 LRAA regulatory limits) when establishing their compliance targets under the proposed rule
- the forecast of compliance choices to be made by utilities

- unit treatment costs for both capital outlays and operations and maintenance (O&M), for each of the viable compliance options, for each utility size and source water
- site-specific retrofit challenges and associated costs
- the allocation of costs of some utility compliance efforts between Stage 1 and Stage 2
- the potential unintended consequences of chloramine conversion (e.g., reduced valve elasticity and life)
- the cost impacts associated with "guidance" elements that in reality become de facto regulatory provisions (e.g., for significant excursions).

Given these and other issues, EPA's stated confidence intervals seem unrealistic. If EPA is going to generate confidence intervals (which it should), then it must consider and reveal clearly what level of realism and precision applies at each stage of the cost analysis, and must track how variabilities and uncertainties can compound across stages of the assessment. The cost estimates provided by EPA, despite their presentation of confidence bounds, do not realistically or usefully portray the expected level or uncertainty range for potential compliance costs.

Stage 2 costs may no longer be a modest increment from Stage 1

The process used to estimate national compliance costs for the Stage 2 D/DBP rule followed a multistep procedure that relied on marginal impact assessments from a predicted Stage 1 baseline to the estimated impact of the Stage 2 construct. Apart from the likely underprediction of impacts on utilities of the Stage 2 compliance assessment using the ICR inputs to the SWAT model, the prediction of compliance costs are likely to underrepresent the true cost due to the use of the marginal cost differential from Stage 1 to Stage 2 technology applications. In essence, this calculation assumes nationally that the only technologies to be costed for Stage 2 compliance are those shifts in technology application from Stage 1 to Stage 2 compliance. For example, if 5% of systems were predicted to install ozone to comply for Stage 1 and 7% of systems were predicted to install ozone for Stage 2 compliance, then the compliance cost for Stage 2 was estimated to include only 2% of systems installing ozone treatment.

At the time of the FACA negotiations, this method of estimating compliance cost was deemed pragmatically reasonable. The time constraints and the number of alternative Stage 2 regulatory frameworks required the use of quick, order of magnitude analyses that could inform the FACA Advisory Committee of the relative differences in national compliance costs associated with the various options. AWWA agreed then and continues to agree that this method of estimating national compliance costs was reasonable for the context of the FACA negotiations. However, AWWA believes that for the proposed rule, the assumptions and methods used to estimate

national compliance costs should have been evaluated further and refined to reflect a more mature assessment of Stage 2 compliance implications in the intervening years.

At the time of the FACA negotiations, implementation of Stage 2 was anticipated to closely follow the implementation of the Stage 1 rule. This proximity in timing contributed to the acceptance of the marginal costing approach because many utilities were expected to actually implement only a single strategy to achieve simultaneous compliance with both Stage 1 and Stage 2. The reality, however, has changed significantly because Stage 2 implementation has been delayed by several years, and the expectation that utilities can implement a single, cohesive strategy has become unrealistic. Therefore, investments have been made by many utilities for Stage 1 compliance that are not relevant to the investment needed for Stage 2 compliance. As such, the true cost for Stage 2 compliance cannot be solely a differential cost from Stage 1 compliance.

As an illustration, consider the following example:

> A plant determined by SWAT to exceed the Stage 1 rule may achieve compliance by switching to chloramines. That same plant may have to install GAC to achieve Stage 2 compliance, but does not require chloramines. In the EPA cost estimates, the cost of compliance for Stage 2 for this plant is the cost of GAC minus the cost of ammonia addition. In reality, this plant incurs both the cost of GAC and ammonia addition since compliance with Stage 1 is sufficiently removed from Stage 2 that incurring the expense of GAC in advance of the requirement is undesirable. The initial conversion to chloramines therefore must be considered a sunk cost. Thus, the Stage 2 cost should be that of GAC implementation in its entirety.

Since the FACA deliberations, sufficient time has been available for re-evaluating the costing method used to estimate Stage 2 compliance impacts. At a minimum, the input data to the SWAT model (AUX 8 Database) could have been revised to reflect the technology selected for each plant to meet the Stage 1 baseline and evaluate the technology forecast resulting from this baseline analysis for Stage 2 costing.

A preliminary review of the model predictions for the post-Stage 1 baseline condition found that 32 of the 273 modeled plants would exceed the Stage 2 LRAA limits of 80/60 with a 20% safety factor. This translates to an estimated 11.7% of large surface water systems being affected by Stage 2. When the marginal estimates of Stage 1 and Stage 2 noncompliance were used, the predicted impact on these same systems was only 2.5%. Therefore, using the approach applied during the FACA negotiations to predict the impact of Stage 2 on large surface water systems for the final compliance assessment substantially underpredicts the impacts (2.5% versus 11.7% of systems impacted).

Closer examination of the 32 plants identified as exceeding the Stage 2 limits using the modeled outcome for the post-Stage 1 baseline indicated the following rate of technology application:

- conversion to chloramines: 72%
- use of conventional treatment methods (enhanced or turbo coagulation, moving chlorine application point): 57%
- UV disinfection: 13%
- chlorine dioxide: 6%
- ozone: 9%
- GAC10: 9%
- GAC20: 3%.

Using the unit cost curves developed by EPA to support the Stage 2 cost estimates, the application of the above technologies for the 11.7% impacted large surface water systems results in a total annualized national compliance cost of $45 million per year. The capital costs nationally were estimated to be $280 million with annual operational and maintenance costs of $27 million. Using the marginal analysis of Stage 1 baseline and Stage 2 estimated impacts, substantially lower national compliance costs were estimated (basically, a negligible cost implication was found for Stage 2 implementation when UV was included as a compliance option). This additional $45 million per year nearly doubles EPA's current estimates of $59 million to $65 million.

Per household cost estimates and affordability

EPA estimates that the cost per household of Stage 2 compliance will in general be fairly modest, at $0.51 and $8.52 per household per year when averaged across all systems and only those systems requiring additional treatment, respectively. However, the mean household level increase would be $43.78 per year in "small systems" needing to add treatment, and nearly 9% of households in this category would be expected to face an increase in their annual water bill of over $120 (U.S. EPA, 2003a, Exhibit ES.8).

A critically important consideration related to interpreting the above EPA results is that the EA designates any CWS of 10,000 or less served as a "small system." The reported per household costs thus reflect a very large range in system sizes (e.g., from 25 served to 10,000 served). This reflects a range that varies by a factor of 400 (10,000/25) in population served within the "small system" size category definition within which EPA reports its key findings. Economies of scale

can be appreciable over this size range, and the aggregated results depicted by EPA are likely to mask much higher per household cost burdens through the smaller end of the CWS size spectrum.

For example, in one location in the EA (U.S. EPA, 2003a, Exhibit 8.4c), the per household costs are sufficiently disaggregated that some insight can be gained on the differences in per household costs within the EPA's overly broad designation of "small" systems. Exhibit 8.4c reveals that the mean annual per household increase is shown to be $185 for systems of under 500 persons served (compared to a mean of $33 for households served by systems in the 3,300 to 10,000 served end of the range). Further, 5% of the households in the under 500 served category will face annual increased costs projected by EPA to be over $409 (U.S. EPA, 2003a, Exhibit 8.4c).

In addition, for customers in systems of under 100 served, the mean per household costs probably will be considerably higher than $185 per year (as predicted by EPA for systems of up to 500 served), and a notable percentage (greater than 5%) of such households may face annual increased costs above $500. EPA does not reveal the costs on a CWS under 100 served basis, but even the limited disaggregation found in the EA's Exhibit 8.4c reveals how much important information is masked under the inappropriate aggregation of all CWS of 10,000 served or less within the "small system" category. Finally, the cumulative impact on affordability may be significant in very small systems that need to comply with Stage 2 plus one or more other rulemakings (e.g., arsenic, radionuclides).

The key point here is that EPA should provide more disaggregated cost results. This is important because there are key equity and affordability implications that are masked (hopefully, unintentionally) by EPA under the approach the Agency uses in the EA to portray its cost and affordability findings. Since the costs are developed using the typical nine system size category scheme EPA has usually employed in the past, the results are all generated by (and available to) the Agency at that level of disaggregation. Merging the five size categories into one overly broad "small system" category of 10,000 served or less is an extra step made by the Agency, and one that obscures rather than informs public review and Agency decision-making.

7.2 Comparison of Benefits to Costs

As AWWA, OMB, NDWAC, SAB, and other organizations have clearly noted on many occasions (e.g., U.S. EPA, Science Advisory Board, 2000a), it is vital that EPA provide suitable and informative comparisons of benefits to costs. The suitable framework is to reveal incremental benefits, incremental costs, and incremental net benefits (i.e., incremental benefits minus incremental costs) for each relevant regulatory alternative. The increments should start with the suitable regulatory baseline and move to increasingly stringent (costly) alternatives.

Ideally, the preferred regulatory option would be identified where the last incremental net benefits is still positive (i.e., just before incremental net benefits turn negative).

Further, it is important that the incremental net benefits be provided not only at the national aggregate level but also according to informative system size categories (U.S. EPA, Science Advisory Board, 2000a). That is, incremental net benefits should be reported for each of the nine CWS size categories EPA usually uses to build its cost and benefit estimates. Additional disaggregation is also worth portraying where important distinctions are reflected in costs, benefits, or both (e.g., other relevant levels of disaggregation might include separating results for groundwater systems from those derived for surface water systems).

Regrettably, EPA fails to provide any meaningful benefit-cost comparisons along the lines described above. All benefits estimates are shown only at a national aggregate level, with no disaggregation by system size. Costs are at times disaggregated, but only across very broad (and therefore meaningless) categories based on whether they serve 10,000 or fewer persons or over 10,000 persons. Further, there is no presentation evident in the EA of incremental net benefits. Overall, the EA is severely disappointing in this regard, and completely at odds with both best practices and recommendations issued in association with past critiques (including the independent reviews of the arsenic EA by the SAB and a NDWAC working group).

7.3 Conclusions

We have not conducted a detailed review of EPA's cost estimates for the proposed Stage 2 rule. Nonetheless, it seems unlikely that the information provided regarding confidence intervals is realistic. Many key uncertainties and variabilities appear to have been either ignored or understated to a considerable degree in order to generate a very narrow (+/- 8%) 90% confidence range. We also believe the Agency may have significantly underestimated costs, because of the likelihood of higher occurrence (e.g., SWAT projections versus utility-specific IDSE evaluations), and because the incremental nature of the rulemaking relative to Stage 1 has been altered by the longer than anticipated intervening time period.

A critical disappointment with the Agency's cost and affordability analyses, and with the benefit-cost comparisons portrayed in the EA, is the lack of meaningful disaggregation according to system size. The lumping by EPA of size categories serves only to mask and obscure important information regarding the equity and efficiency implications of the proposed rule. This is a serious flaw and a considerable disservice to the public, stakeholders, and decision-makers.

Finally, the Agency's approach to comparing benefits to costs is seriously flawed, because there are no incremental net benefits revealed, nor are benefit-cost findings presented on a system size basis. The Agency's reliance on only total (rather than incremental) benefits and costs, and its

gross aggregation of all findings to the national level (rather than revealing outcomes according to system size categories as well), is very troubling and at odds with good practices and the recommendations of expert review panels, including SAB and NDWAC. The Agency needs to do much better, and it will take only a modest effort to do so (if it musters the will).

References

Bick, R.L., J. Madden, K.B. Heller, and A. Toofanian. 1999. Recurrent Miscarriage: Causes, Evaluation, and Treatment. Medscape General Medicine 1(3). Available at http://www.medscape.com/viewarticle/408885

Bove, F.J., M.C. Fulcomer, J.B. Klotz, J. Esmart, E.M. Dufficy, and J.E. Savrin. 1995. Public drinking water contamination and birth outcomes. *American Journal of Epidemiology* 141(9).

CDC. 2003. Prevalence of Current Cigarette Smoking among Adults and Changes in Prevalence of Current and Some Day Smoking — United States, 1996 — 2001. Centers for Disease Control. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5214a2.htm. Accessed November 19, 2003.

Crawford-Brown, D. 2003. Review of Health Effects Issues Associated with the Benefits Analysis of the EPA for DBP and Microbial Risk. Prepared for AWWA. October.

Dana Farber Cancer Institute. 2003. Cancer Treatment Information. Internet: http://www.dana-farber.org/can/cancer/. Accessed October 17, 2003.

Dodds, L., W.D. King, C. Wolcott, and J. Pole. 1999. Trihalomethanes in public water supplies and adverse birth outcomes. *Epidemiology* 10(3):233-237.

Hansen, K. 1998. Seasonality of Moves and Duration of Residence. U.S. Census Bureau Report #P70-66.

King, W.D. and L.D. Marrett. 1996. Case-control study of bladder cancer and chlorination by-products in treated water (Ontario, Canada). *Cancer Causes Control* 7.

King, W.D., L. Dodds, and A.C. Allen. 2000. Relation between stillbirth and specific chlorination by-products in public water supplies. *Environmental Health Perspectives* 108(9).

Li, D-K., L. Liu, and R. Odouli. 2003. Exposure to non-steroidal anti-inflammatory drugs during pregnancy and risk of miscarriage: population based cohort study. *British Medical Journal* 327(August).

Magat, W.A., W.K. Viscusi, and J. Huber. 1996. A reference lottery metric for valuing health. *Management Science* 42(8):1118-1130.

Malcolm Pirnie Inc. 1999. Surface Water Analytical Tool (SWAT): Program Description and Assumptions/Specifications. U.S. Environmental Protection Agency, Office of Groundwater and Drinking Water, Washington, DC.

March of Dimes. 2003. Pregnancy & Newborn Health Education Center: Eating for Two. http://www.modimes.org/pnhec/159-823.asp. Accessed December 9, 2003.

NCI. 1998. *What You Need to Know About Bladder Cancer.* National Cancer Institute. http://cancernet.nci.nih.gov/wyntk_pubs/bladder.htm. Downloaded July 21, 2000.

NCI. 2000. Surveillance, Epidemiology, and End Results (SEERS). National Cancer Institute. http://seer.cancer.gov/csr/1973_1997/index.html. Accessed December 16, 2003.

NCI. 2003. Contents of the SEER Cancer Statistics Review, 1975-2000. http://www.seer.cancer.gov/csr/1975_2000/sections.html (links accessed: SEER Incidence, US Mortality, Relative Survival Rates: All Races; Age Distribution Diagnosis and Death; Median Age at Diagnosis and Death; Years of Life Lost; 5-Year Relative Survival Rates). Accessed October 17, 2003.

Nielsen, G.L., H.T. Sorensen, H. Larsen, and I. Peterson. 2001. Risk of adverse birth outcome and miscarriage in pregnant users of non-steroidal anti-inflammatory drugs: Population based observational study and case-control study. *British Medical Journal* 322:266-270.

NRC. 1999. Risk Assessment of Radon in Drinking Water. National Research Council. National Academies Press, Washington, D.C. The number was based on the following study: Ershow, A.G. and F.P. Cantor. 1989. Total Water and Tapwater Intake in the U.S.: Population-based estimates. Bethesda, MD: Federation of American Societies for Experimental Biology, Life Sciences Research Office.

OMB. 2002. *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*; Notice; Republication, 67 Fed. Reg. 8452-8460 (February 22). Office of Management and Budget. Available at http://www.whitehouse.gov/omb/fedreg/reproducible2.pdf.

Raucher. R. et al. 2001. Identifying Geographic Regions for Potential Future Epidemiology Research on Disinfection Byproducts. Prepared for the Awwa Research Foundation, U.S. EPA, and the Microbials/Disinfection By-Products Research Council. Denver, CO.

Raucher, R.S., M.L. Hagenstad, J. Cotruvo, R. Narasimhan, K. Martin, H. Arora, R. Regunathan, J.A. Drago, and F. Pontius. Forthcoming. Conventional and Unconventional Approaches to Water Service Provision. Awwarf Project #2761. Awwa Research Foundation, Awwa Research Foundation, and American Water Works Association.

Roseberry, A.M. and D.E. Burmaster. 1992. Lognormal Distributions for Water Intake by Children and Adults. *Risk Analysis*, 12(1):99-104.

U.S. EPA. 2000a. Guidelines for Preparing Economic Analyses. EPA 240-R-00-003. U.S. Environmental Protection Agency. September.

U.S. EPA. 2000b. Arsenic in Drinking Water Rule Economic Analysis. EPA 815-R-00-026. U.S. Environmental Protection Agency, Office of Water, Washington, DC. December.

U.S. EPA. 2002. *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency* (EPA/260R-02-008, October). Available at http://www.epa.gov/oei/qualityguidelines/-EPA_OEI_IQG_FINAL_10-2002.pdf.

U.S. EPA. 2003a. Economic Analysis for the Proposed Stage 2 DBPR (Disinfection By-Products Rule). EPA 815-D-03-001. U.S. Environmental Protection Agency.

U.S. EPA. 2003b. National Primary Drinking Water Regulations: Stage 2 Disinfections and Disinfections Byproducts Rule; National Primary and Secondary Drinking Water Regulations: Approval of Analytical Methods for Chemical Contaminants; Proposed Rule. 40 CFR Parts 141, 142, and 143. August 18, 2003.

U.S. EPA. 2003c. Analysis and Findings of the Gallup Organization's Drinking Water Customer Satisfaction Survey. U.S. Environmental Protection Agency, Office of Groundwater and Drinking Water. August 6.

U.S. EPA. 2003d. Economic Analysis for the Proposed Long Term 2 Enhanced Surface Water Treatment Rule. prepared for the U.S. Environmental Protection Agency by the Cadmus Group. June.

U.S. EPA, Science Advisory Board (Drinking Water Committee). 2000a. An SAB Report on Certain Elements of the Proposed Arsenic Drinking Water Regulation. Executive Committee Review Draft of September 6.

U.S. EPA, Science Advisory Board. 2000b. An SAB Report on EPA's White Paper *Valuing the Benefits of Fatal Cancer Risk Reduction.* Internet: http://www.epa.gov/sab/pdf/eeacf013.pdf. Accessed November 10, 2003.

Ventura, S.J., W.D. Mosher, S.C. Curtin, J.C. Abma, and S. Henshaw. 2000. Trends in pregnancies and pregnancy rates by outcome: Estimates for the United States, 1976-96. National Center for Health Statistics. *Vital Health Stat.* 21(56).

Waller, K., S.H. Swan, G. DeLorenze, and B. Hopkins. 1998. Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology* 9(2):134-140.

Waller, K., S.H. Swan, G.C. Windham, L.Fenster. 2001. Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. *Journal of Exposure Analysis and Environmental Epidemiology*. 11(6):522-531.

WQA. 1999. 1999 National Consumer Water Quality Survey. Water Quality Association. Lisle, Ill.: WQA.7.

Appendix: Problems with EPA's Use of ICR Data in its Fetal Loss Calculations

EPA's use of ICR data to quantify fetal losses avoided is not transparent or reproducible. The ICR query language included in Appendix B of the *Occurrence Assessment of the Stage 2 DBP Rule* is flawed and does not match results reported in either the *Occurrence Assessment for the Stage 2 DBP Rule* or the *Economic Analysis for the Stage 2 DBPR Proposal*. To analyze the results presented in the *Occurrence Assessment* and *Economic Analysis*, independent queries had to be developed. In doing so, AWWA concluded that the key tables related to fetal loss assessments (notably Exhibit 5.13 of the *Economic Analysis for the Stage 2 DBPR Proposal*) were correct for the TTHM study level of 80 µg/L. One exception was the number of locations with peaks above the TTHM study level of 80 µg/L under pre-Stage 2 conditions; Exhibit 5.13 identified 73 locations while the independent queries identified 72 locations. In addition, EPA queries repeatedly identify an erroneous result contained within the ICR database. The ICR database reports a TTHM concentration of 300 µg/L for Plant 753 at the AVG1 location during sample quarter 3; the result should be recorded as 30.0 µg/L.

In addition to transparency and reproducibility issues with ICR data, EPA's estimate of 1,100 to 4,700 fetal losses avoided is incorrect due to a simple typographic error. Step 1 of the estimate (*Economic Analysis for the Stage 2 DBPR Proposal, Appendix G, page G-6*) shows the baseline number of fetal losses attributable to exposure to peak DBPs as

$$1.7\% \text{ PAR} \times 983{,}000 = 16{,}711.$$

In Step 3, the previously calculated result is misrepresented as 16,911.

Following the results in EPA's Exhibit 5.13, the Stage 2 DBP rule implementation will result in the reduction of TTHM results above a study level of 80 µg/L from 73 locations to 19 locations. Since the independent queries developed by AWWA identified 72 locations with TTHM results above 80 µg/L rather than 73 as reported in Exhibit 5.13, 72 locations will be assumed for the purpose of this discussion. Therefore, 53 locations that exhibited TTHM results above 80 µg/L under pre-Stage 2 conditions were eliminated under post-Stage 2 conditions. From these 53 locations, 37 plants are represented from 30 utilities.

It is also important to understand the species mixture of the 53 quarterly ICR TTHM results above a study level of 80 µg/L that are eliminated under post-Stage 2 conditions. Exhibit A.1 displays the species contributions for the 53 TTHM results.

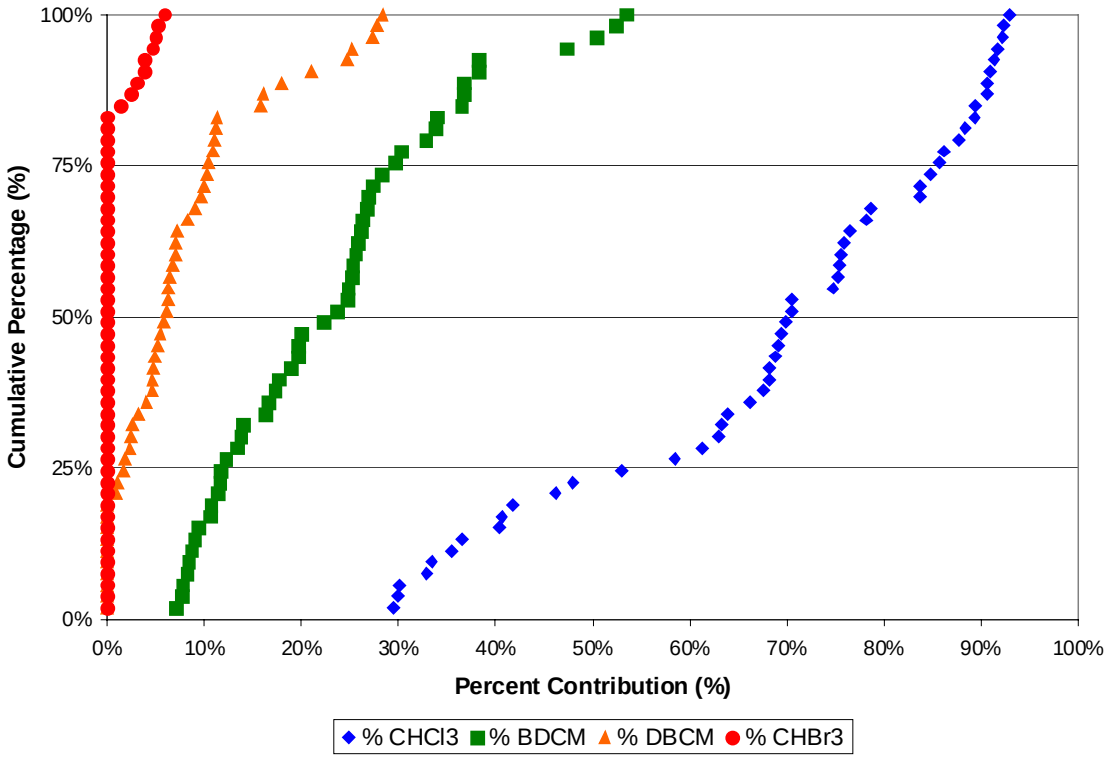


Exhibit A.1. Species contributions for TTHM results > 80 µg/L eliminated under post-Stage 2 conditions.

As seen in Exhibit A.1, the 53 TTHM results are dominated by chloroform ($CHCl_3$). In fact, the median percent contribution of chloroform to the 53 TTHM results is 71%. On the other hand, the median percent contributions of DBCM and bromoform to the 53 TTHM results are 6% and 0%, respectively.

A.1 ICR Data Analysis Transparency Issues

Section 1.4.8 of the *Occurrence Assessment for the Stage 2 DBP Rule* outlines the methods and assumptions for ICR data analysis used throughout the *Occurrence Assessment for the Stage 2 DBP Rule* and the *Economic Analysis for the Stage 2 DBPR Proposal*. Appendix B of the *Occurrence Assessment*, Section B.3, includes the query language used for plant screening. AWWA commends the inclusion of the query language to support transparency of the data analysis completed for the proposed rule. However, the query language contained within

Appendix B of the *Occurrence Assessment*, contains errors and does not follow the same assumptions and methodology described in Section 1.4.8 of the *Occurrence Assessment*.

Exhibit A.2 describes the relationships between all the queries included in Appendix B. Errors in the query language are pointed out in bold.

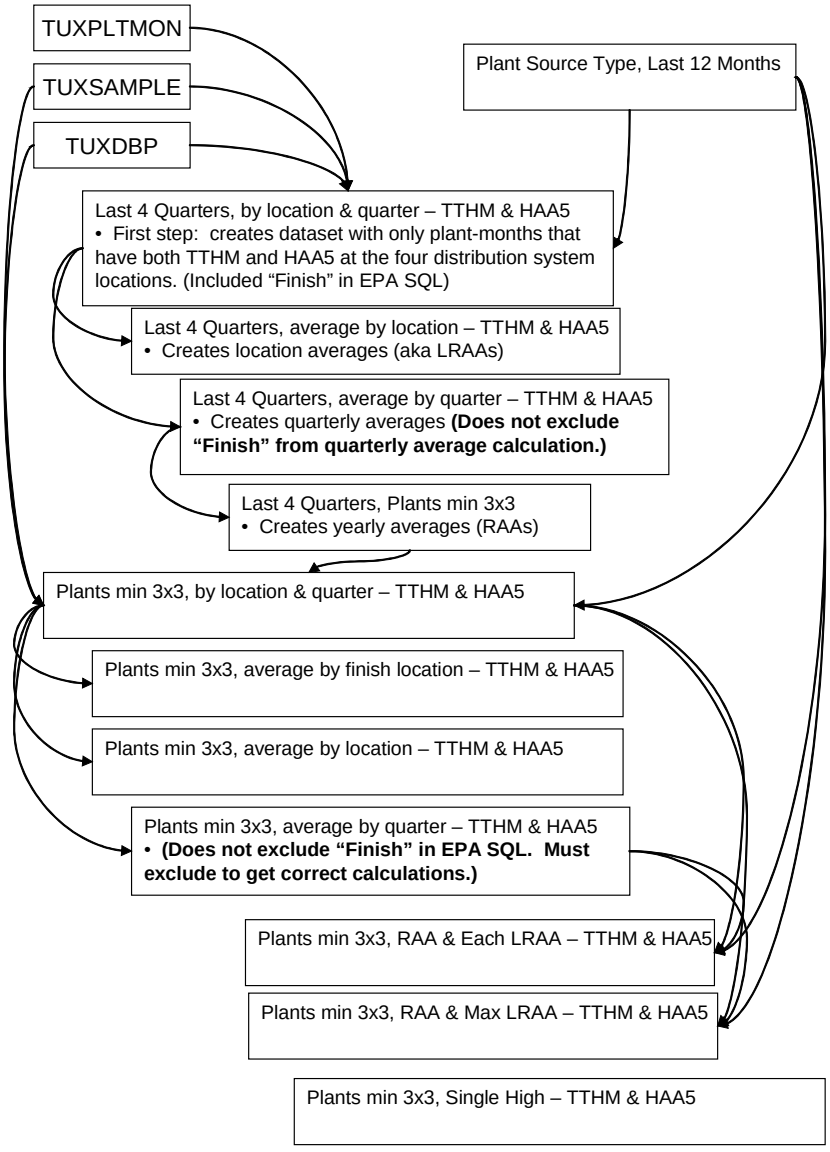


Exhibit A.2. EPA query language diagram.

As indicated in Exhibit A.2, the inclusion of Finish location DBP results in quarterly average calculations in the "Last 4 Quarters, average by location — TTHM & HAA5" and "Plants min 3x3, average by quarter — TTHM & HAA5" queries results in errors of subsequent RAA and LRAA calculations. While the query language contains these errors, the results included in the *Occurrence Assessment* and *Economic Analysis* that reference these queries are in fact correct. To ensure transparency of the rule's supporting data analysis, the corrected query language should be included in Appendix B of the *Occurrence Assessment for the Stage 2 DBP Rule*.

Appendix B.
Critique of the Economic Analysis for the Proposed Long-Term 2 Enhanced Surface Water Treatment Rule, Final Report, Stratus Consulting

Stratus Consulting

Critique of the Economic Analysis for the Proposed Long-Term 2 Enhanced Surface Water Treatment Rule

Final Report

Prepared for:

AWWA
1401 New York Ave., NW
Suite 640
Washington, DC 20005

Prepared by:

Robert Raucher, PhD
David Mills
Lauraine Chestnut
Stratus Consulting Inc.
PO Box 4059
Boulder, CO 80306-4059
(303) 381-8000

January 8, 2004
SC10364

Critique of the Economic Analysis for the Proposed Long-Term 2 Enhanced Surface Water Treatment Rule

Final Report

Prepared for:

AWWA
1401 New York Ave., NW
Suite 640
Washington, DC 20005

Prepared by:

Robert Raucher, PhD
David Mills
Lauraine Chestnut
Stratus Consulting Inc.
PO Box 4059
Boulder, CO 80306-4059
(303) 381-8000

January 8, 2004
SC10364

Contents

Executive Summary

This report provides a critique of the U.S. Environmental Protection Agency (EPA) Economic Analysis (EA) of the proposed Long-Term 2 Enhanced Surface Water Treatment Rule (LT2). The EA provides a benefit-cost analysis of the rule. Our review of the EA methods, data, results, and presentation by EPA reveals some good efforts by the Agency, but also several critical areas of concern that require extensive improvement.

Our major observations and findings with respect to the EA include the following:

1. Overall, we believe EPA has considerably overstated the occurrence and risks associated with endemic levels of *Cryptosporidium* in finished waters, and thus the Agency overstates the benefits of the proposed rule to a considerable degree. The costs of the rule may also be overstated to some degree.

2. The ICRSS data indicate a much smaller percentage of systems will end up in bins 3 and 4 under the proposed rule than do the analyses based on the ICR data, implying that the net benefits (benefits minus costs) of the proposed rule may be 20% of the high end estimates shown by EPA (all else equal). The ICRSS data are better predictors than the ICR data are of what the impact of the rule will be as proposed.[1]

3. EPA applies a Bayesian interpretation to the ICR and ICRSS data that is suspect and driven by unsubstantiated and perhaps extreme assumptions. For example, EPA imposes an assumption that only 1 out of every 1,000 "zeroes" observed in the database is truly a zero. The Agency is thus estimating occurrence and risk based on a presumption that 999 out of every 1,000 observed zeroes in the database are instead one oocyst or more.

4. EPA's exposure assessment is based on considerably overestimated levels of direct ingestion of CWS-provided waters. Relevant exposures (and, hence, risks) may be overstated by a factor of 2 or 3 when direct ingestion rates for CWS waters, and increased bottled water use, are properly considered.

1. The ICRSS data are more indicative of what the rule's impacts will be because they (1) probably are more accurate than the ICR data (ICRSS results are based on Method 1622/1623 with higher recovery rates than the IFA method applied in the ICR data), and (2) reflect the method (1622/1623) that utilities will apply in their compliance monitoring.

5. The infectivity dose-response relationship applied by EPA is subject to considerable uncertainty and probably overstates the risk associated with exposures to an infectious oocyst by a significant degree.

    a. The underlying clinical studies use extremely high doses relative to oocyst levels in finished waters (levels of oocysts ingested of 23,000 to 2.3 billion times higher than now found in finished waters) and rely on extremely small numbers of subjects and strains (between 14 and 29 subjects, for each of only three strains).

    b. The results of the clinical studies are interpreted liberally, based on a "presumed infection" approach that assumes that any subject with symptoms has cryptosporidiosis, even when several of the symptomatic subjects had no documented infection (e.g., positive oocyst shedding). EPA's risk estimates are overstated to the extent that reported symptoms could be attributable to causes other than cryptosporidiosis.

    c. The results of the clinical studies were interpreted via complex statistical models that are driven by — and highly sensitive to — unsubstantiated assumptions. While the modeling approaches used by EPA in the EA were suggested by the SAB, the obscurity of the presentation and the sensitivity of the results to the model assumptions (e.g., increasing a key estimated mean risk parameter by a factor of 4 or 5 over the level found in the peer reviewed published literature) reveal the need for more transparency, continued scientific discourse, and greater use of sensitivity analyses in portraying the possible risk levels.

6. The extent by which EPA's risk model overstates risks can be viewed, in part, by comparing the Agency's estimated number of waterborne cases of cryptosporidiosis at the pre-LT2 baseline to its estimated reduction in cases due to the proposed LT2 rule:

    a. EPA estimates that the pre-LT2 baseline (i.e., post IESWTR) is between 60,000 and 111,000 cases per year.

    b. The Agency's risk model used for the LT2 rule benefit-cost analysis predicts 256,000 to over 1,000,000 cases per year will be avoided because of the rule as proposed.

    c. Therefore, EPA estimates a reduction in cases that is up to 9+ times higher than the number of cases it has stated exist at baseline.

7. EPA needs to explore the soundness and implications of its questionable assumption that the risk of illness (as well as severity and duration of illness) is independent of dose. The morbidity assessment — used to project the number, severity, and duration of illnesses

due to a possible infection – is based exclusively on results from the Milwaukee outbreak of 1993, where oocyst levels were much higher, exposure durations much longer, and opportunities for secondary spread and exposure more pervasive than anticipated under the endemic low dose exposure context addressed by the proposed rule.

8. EPA's use of an "enhanced" cost-of-illness (COI) approach to value avoided cases of nonfatal cryptosporidiosis is highly problematic. The approach is a significant departure from standard economics practice, does not appear to have been subjected to peer review, and yields results that seem implausible and unrealistic compared to other well-established risk valuation benchmarks.

9. EPA's presentation of regulatory costs and benefits is overly aggregated, and fails to reveal how affordability and net benefits vary across system size categories or across other relevant program elements in the proposed rule (e.g., reservoir covering, filtered versus unfiltered systems).

1. Introduction

This report provides a critique of the Economic Analysis (EA) (U.S. EPA, 2003a), developed by the U.S. Environmental Protection Agency (EPA), in support of its proposed rulemaking for the Long-Term 2 Enhanced Surface Water Treatment Rule (referred to hereafter as LT2), published in the *Federal Register* on August 11, 2003 (U.S. EPA, 2003b). In essence, the EA is a benefit-cost analysis, and as such it contains many critical elements that support the rationale for the rulemaking and that are used to predict the rule's expected impacts on public health and water costs. The EA includes issues of occurrence, exposure, dose-response relationships, quantified risk characterization, and the valuation of health risk reductions. Issues related to compliance technologies, and their performance and costs, are also integral to the EA.

1.1 Objectives

The objective of this critique is to focus on key aspects of the Agency's empirical analysis, and highlight where there appear to be actual or potential errors or omissions of consequence in the EA. Our primary concerns are with the quality, transparency, replicability, objectivity, clarity, and reliability of the data and methods used to develop and present estimates of the benefits and costs of the proposed rule. Our focus is on the proper use and interpretation of sound methods and data, the presentation of the approach and outcomes in a clear and unbiased manner, and the drawing of suitable inferences and policy interpretations given the uncertainties and other limitations inherent in the data and analyses. The overall objective is to help EPA identify and remedy potential errors and limitations in its own EA, and thereby help EPA generate more reliable, accurate, and policy-useful benefit-cost analyses for the final LT2 and other rules in the future.

No attempts are made here to develop or report independent estimates of the proposed rule's benefits, costs, or incremental net benefits. Instead, we focus on identifying aspects of EPA's analysis that appear to require correction or, at a minimum, further investigation and documentation. The overall goal is to help identify areas in which improvements can and should be made, so that EPA's EA of the final LT2 rule — and of other future rulemakings — can embody improved quality, integrity, and meaning for the important public health protection policies the Agency pursues in accordance with the mandates of the Safe Drinking Water Act (SDWA). We hope and expect that EPA will embody changes in the EA supporting the final LT2 Rule to address the key points raised in this report, and to represent the most accurate benefits and costs to the public as possible.

1.2 Overview of Critique

Our critique of the EPA EA and associated Agency-provided documentation focuses on the need to find a better balance and sense of prioritization in its analyses and presentation. For example:

- EPA is to be commended for provided a vast quantity of detail and back-up documentation. On the other hand, there is too much material to effectively review within the comment period. More important, EPA is inconsistent in that its documentation is sometimes lacking enough information on some core issues, while offering mountains of data on other matters of sometimes lesser significance. In general, more balance and prioritization is needed in terms of detail and documentation.

- EPA is to be commended for attempting to quantify benefits to the greatest degree feasible, so that benefits can be compared to costs. On the other hand, EPA has in places stretched credibility and the bounds of "good science" in its efforts to generate estimates. "Good science" must always be the guiding principle, and empirical estimates based on interpretations or assumptions that have strayed from that principle are potentially misleading and a disservice to the public, stakeholders, and decision-makers.

- EPA has made some good strides forward by providing considerable discussion and sophisticated numeric evaluation of several of the uncertainties and variabilities (e.g., using Monte Carlo simulations) applicable to some aspects of the analysis. On the other hand, EPA neglects to detail, justify, or fully explore some of the most fundamental of its assumptions. In the face of these core uncertainties, transparent sensitivity analyses are essential for evaluating the impact of core assumptions at key junctures of the analysis. EPA needs to find a better balance by using more fundamental, informative, and simple analyses of core components rather than using more sophisticated approaches for some lesser aspects of its analysis.

Hence, at the core of our critique is the message that EPA needs to take better stock of its analyses, determine what components are most critical in terms of driving the benefit or cost estimates, and focus its attention (and that of the reviewers) on those aspects. Models, analytic tools, and documentation should be presented in a way that sheds light rather than obfuscates and overwhelms attempts at good faith public review.

1.3 Outline of this EA Critique

Following this introductory chapter, the following analysis is provided:

- Chapter 2 provides a schematic overview of the EPA benefits analysis for the proposed LT2 rule, along with a summary of the key concerns identified by our review.
- Chapter 3 explores issues with the first set of steps in the analysis, related to the use and interpretation of occurrence data.
- Chapter 4 explores exposure issues and their impacts on the analysis.
- Chapter 5 discusses the dose-response elements of the analysis, notably the infectivity and morbidity assessments for exposures to *Cryptosporidium*.
- Chapter 6 critiques the valuation of risk reductions, notably the enhanced cost of illness (COI) approaches applied by the Agency.
- Chapter 7 evaluates the Agency's cost analyses, and the comparison of benefits to costs.

1.4 Summary Conclusion

The EA and associated EPA-provided documents offer extensive detail, information, and background material. It is evident that EPA has explored many aspects of the analysis and has provided a considerable body of documentation. However, some key elements of EPA's work are not sufficiently documented or detailed for an effective review, and this has hampered our evaluation. Overall, the Agency needs to find a more suitable balance.

In some very critical elements of the EA, the Agency makes strong assumptions that can have a significant impact on the final results, yet the Agency does not always clearly articulate what assumptions it is making or it appears to take a one-sided view of the uncertainties and data limitations to derive its interpretation. More important, in these instances of strong assumptions, EPA has in several instances failed to offer sensitivity analyses based on equally or more plausible alternative assumptions. EPA should follow the basic tenants of its own *Guidelines for Economic Analyses* and (a) be explicit regarding its core assumptions, (b) document the basis for those assumptions, and (c) develop some useful sensitivity analyses to evaluate and convey the impact of these core uncertainties on the outcomes of the analysis.

2. Overview of Benefits Analysis

To best explain the nature of the EPA analysis, a simplified flow diagram is shown in Exhibit 2.1 to identify and describe the key steps of the benefits assessment. The same underlying logic and data also underlie the basis for the cost analysis (although the path of the cost analysis naturally differs from that of the benefits after the initial occurrence assessment).

Along with each critical step in the analysis, Exhibit 2.1 displays a summary of some of the key issues and concerns regarding each of these steps. Additional detail is then provided in each subsequent chapter of this report.

It is apparent when examining the issues identified in Exhibit 2.1 that there are many ways in which the risk reduction benefits developed by EPA may be overstated. Some components of the cost analysis may be overstated as well, especially with regard to the number of water systems projected to be in higher "bins" (i.e., bins 3 and 4). Therefore, EPA may over-predict the number of systems with relatively high compliance costs (and benefits).

The "big picture" view of the EPA analysis, as afforded by Exhibit 2.1, allows us to glean how much the overall EPA results might be altered if alternative (but equally or more plausible) assumptions and data interpretations were investigated. Compounding the changes at each step as depicted in the figure, the overall estimate of benefits derived by EPA could be an order of magnitude or more larger than a more plausible and likely estimate. Detailed discussions are provided in the chapters that follow.

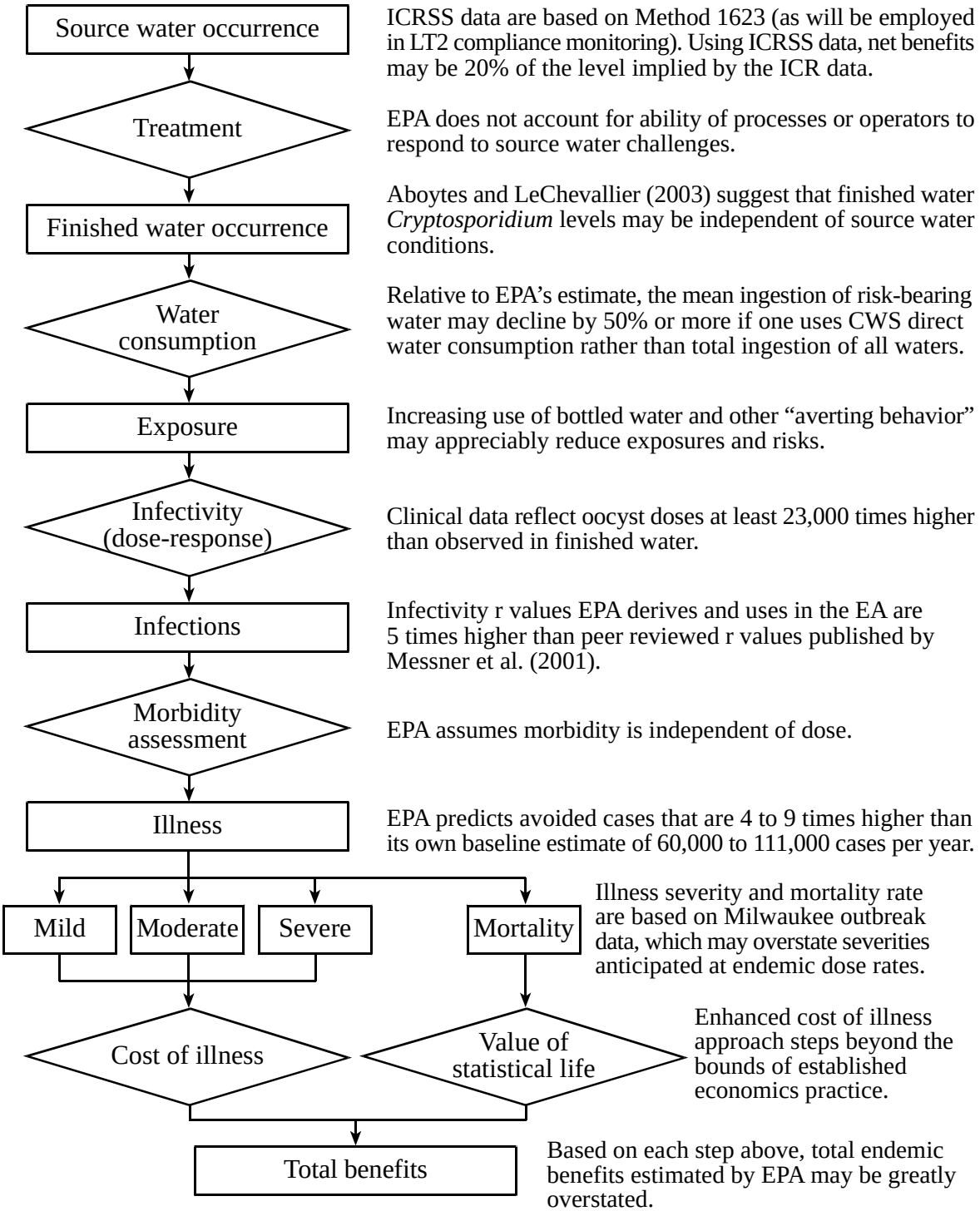


Exhibit 2.1. Benefits analysis process.

3. Occurrence Assessment: Source and Finished Water Characterization

3.1 Source Water Occurrence Data

EPA uses data collected under the Information Collection Rule (ICR), and two subsequent smaller ICR Supplemental Surveys (ICRSS), to assess the presence of *Cryptosporidium* oocysts in source waters.

- The initial ICR reflects data collected from nearly 300 Community Water Systems (CWS) serving more than 100,000 people. Cryptosporidium monitoring was conducted for 350 influent locations (i.e., plants using surface waters or groundwater under the influence of surface waters), including unfiltered systems. The data were collected over the 18 month period from July 1997 through December 1998. The analysis relied on the "ICR method" of *Cryptosporidium* detection and analysis.

- The ICRSS includes data collected from 40 randomly selected volunteer medium-sized CWS serving 10,000 to 100,000 people (ICRSSM), and 40 randomly selected volunteer large systems of over 100,000 served (ICRSSL). The samples were drawn over a 12 month period spanning March 1999 through February 2000. The analysis relied on the improved analytics of EPA method 1622/1623, which provides better recovery rates than the older method used in the ICR.

There are appreciable differences between the results generated by the initial ICR and the follow-up ICRSS. For example, the ICR data assign 7.6% of the systems into either bin 3 or bin 4, whereas the ICRSS assign an average of less than 1.2% of the systems to those bins — a difference factor of over 6.4 (7.6/1.2). The magnitude of this difference raises questions about which set of ICR results are most accurate and reliable (or, whether the apparent discrepancies reflect real differences due to temporal and/or spatial variability in oocyst occurrence levels).

EPA addressed the uncertainty associated with the differences between the ICR and ICRSS results by conducting sensitivity analyses. The Agency presents benefit and cost estimates based on each data set (i.e., there are three sets of benefit and cost results, one based on the ICR, one on the ICRSSL, and one on the ICRSSM). This is an appropriate approach for addressing the uncertainty about which data set might be the most accurate reflection of actual conditions. Ultimately, the estimated net benefits of the preferred regulatory option differ by a factor of 4.5 to more than 5.2, depending on which dataset is used (EA Exhibit ES.8, e.g., 705/135 = 5.22; U.S. EPA, 2003a).

While the use of sensitivity analysis in this manner is fully appropriate, EPA should also provide further discussion of why the results from the ICRSS are so different from the results derived using the ICR. There are several plausible and likely explanations that, if brought to light, would have provided readers and decision makers with some useful context within which to evaluate the widely different sets of results.

In particular, it is likely that the ICRSS data are more reliable than the ICR data because the ICRSS use a much improved analytic method (1622/1623). While the ICR provides results from a larger sample, the ICRSS results are likely to be considerably more accurate. It is very likely that the higher accuracy and reliability of the ICRSS data outweigh the ICR's advantage of having more data. While we have not attempted to develop or perform any statistical analyses of the potential tradeoff, it seems logical that this is an instance of where more accurate data are more valuable and informative than larger quantities of questionable data. In addition, because source water characteristics are not likely to be associated with the size of the population served by a CWS drawing those waters, it may be suitable and appropriate to consider merging the two ICRSS databases (at least for source water characterization purposes), creating a combined dataset of observations from 80 utilities. Finally, because Method 1623 will serve as the basis for LT2 compliance monitoring (bin determination), the ICRSS results are likely to provide a more accurate characterization of the impact of the rule, in terms of compliance efforts, costs, and benefits.

3.2 Bayesian Interpretation of the Source Water Data

The ICR and ICRSS data present some challenges for interpretation for a national occurrence profile, in large measure because of the limitations of the methods for identifying and counting *Cryptosporidium* oocysts. There is considerable uncertainty about the accuracy of presence/absence findings, and of numbers of oocysts derived in the ICR and ICRSS findings (more so with the ICR because of the lower recovery factors with the analytic method).

For example, there is a very high likelihood of false positives because of limits in the immunofluorescent antibody (IFA) methodology used in the ICR. This would erroneously push more systems into higher bins.

Likewise, there is a large percentage of zeros. This is because *Cryptosporidium* may not be present in many source water samples, and/or because of the low recovery rates associated with the analytic methods available (especially for the ICR data). Positive samples were noted for only 5% of the ICR observations (95% were zeros). Given the variability in recovery, this raises the question of how many of the observed findings of "no oocysts detected" are true zeroes. If some fraction of the nondetects actually reflect cases of a true value of one or more oocysts, then this potentially places systems erroneously in bin 1 instead of a higher bin.

To address the uncertainties associated with the data quality, EPA applied Bayesian statistical techniques to the ICR and ICRSS data. Bayesian techniques can be very powerful and effective tools for dealing with uncertainties. However, an integral part of the Bayesian approach is the use of "informed priors" that reflect what knowledge or outcomes the researchers believe to be true (e.g., using probabilities based on known data). The specific priors that are applied in Bayesian applications typically drive the outcomes that are derived from the analysis. In other words, the priors typically determine the outcomes.

In some instances, there are enough good data and other reliable information available that the priors can be set with some degree of confidence and rationale. However, the less that is known *a priori*, the more the priors become *de facto* (and perhaps unsubstantiated) assumptions. This appears to be the case with the *Cryptosporidium* analytic method results from the ICR databases. Because so little is known about the "true" values as opposed to the numbers of oocysts observed from the analytic methods, the Agency's Bayesian approach is in effect driven by the critical assumptions that EPA established as its priors. EPA's occurrence results are thus driven by the presumption that (1) at least 99.9% of the observed zeroes are in fact integer values of 1 or more (i.e., that at most one out of 1,000 observed zeroes are true zeroes and the remaining 999 observed zeroes are assigned a value of 1 or more oocysts); and (2) there are no more than 1% false positives.

Both of these assumptions appear to be extreme and arbitrary. For example, one expert noted that 25% to 50% of the observed zeroes might be true zeroes, in contrast to EPA's assumption of only 0.1%.[1] This is a critical factor, because if the value of Z_i were to increase by over two orders of magnitude (from 0.1% to 25%), this would change EPA's results dramatically (see EA Appendix B, Section B.2.3). It is important to determine if the EPA number is correct on page B-12 (i.e., 0.1%), and it is important that EPA explain why it believes this number is so low. The text says that the model was tested with percentages from 0 to 50%, but that the Agency's experts thought it should be 0.1%. It is essential that EPA help the public understand the sensitivity of the model to the value of true zeros, and to provide clarification of what is meant by true zero in this document.[2] Appendix A provides additional discussion.

1. Personal communication with Jeff Rosen (Perot Systems), relaying the expert opinion he obtained from Jen Clancy (November 2003).

2. On EA page B-12, EPA states that the true zero assumption is "the true proportion of systems with source water that is completely free of the target microbe." This is not our understanding of the true zeros. The question is what percentage of the zeros were likely to be true zeros and not some other number that needed to be estimated.

Likewise, for false positives, EPA analyses apparently were run at only 0% and 1%. False positives could be much higher, and this distinct likelihood must be addressed in sensitivity analyses as well.[3]

Given that EPA's results are driven by assumptions (over which there was considerable disagreement with stakeholder experts) rather than "prior knowledge," it is imperative that the Agency:

1. Identify explicitly and label clearly what assumptions it is using (i.e., identify them as assumptions and clarify that they do not satisfy the intent of "prior knowledge" as generally applied or accepted in Bayesian analysis).

2. Provide a clear and cogent rationale for why those assumptions are indeed reasonable and defensible.

3. Develop and/or accept alternative assumptions that are equally defensible or more plausible.

4. Conduct sensitivity analyses to understand and reveal how the alternative assumptions (priors) affect the outcomes of the analysis.

5. Present the results of the sensitivity analyses in a clear, objective, and informative manner.

6. Use the results of the sensitivity analyses to inform the decision-making process (and document the same).

3.3 Characterization of Finished Waters

From source water quality, EPA needs to take two steps to assess the impacts of the regulation on oocyst levels of finished waters to which consumers are exposed. First, EPA predicts what finished water will be like without the LT2 rule. This establishes the finished water baseline for the proposed rule. Second, EPA predicts how the proposed LT2 rule will reduce *Cryptosporidium* oocysts in post-rule finished water relative to the baseline finished water.

3. For example, EPA should document the effect of using ICR data on total *Cryptosporidium*, versus non-empty oocysts, versus oocysts with internal structure, and justify why the Agency used total *Cryptosporidium* to do the assessment. The original sensitivity analyses run suggests that the total is one log higher than the non-empty option, which is one log higher than the internal structure approach.

Unfortunately, there has been limited finished water analysis of *Cryptosporidium* levels, so results for both of the above steps were projected by EPA using various models and assumptions. Research by Aboytes et al. (2000) is a notable exception and provides some finished water data that EPA uses as a point of comparison to their projected (model-driven) outcomes.

In each step, EPA assumes a mix of treatment choices will be made by utilities, and the Agency applies expected treatment performance information to estimate the occurrence levels of *Cryptosporidium* oocysts in finished water (i.e., predicting changes from estimated finished water conditions pre LT2). The technology selections and performance levels that EPA assigns may be reasonable in general, although a few problems have been noted. For example, chlorine dioxide is not considered a viable choice for 0.5 or better log removals (e.g., as depicted in EA Exhibit F-17), and therefore the 23% of systems that EPA predicts will select this approach should instead be set at 0%. These systems instead should be assigned to turbidity control or UV light in the compliance forecast, and costs adjusted accordingly.

Of greater concern for the EA, however, is that the EPA approach does not reflect the likely ability of treatment processes and plant operators to handle challenges posed by source water conditions. There is likely to be more of a leveling effect on finished water quality across source water conditions than is reflected in the EPA approach. The research by Aboytes et al. (2000) suggests that source water quality may not be a major determinant of finished water outcomes with respect to oocyst levels.[4]

3.4 Conclusions

Estimating source and finished water levels of oocysts is extremely challenging because of the many limitations inherent in the analytic methods that define the results found in the ICR and, to a lesser extent, the ICRSS. Ultimately, the ICRSS findings may be more reliable. The ICRSS findings indicate relatively low baseline oocyst occurrence and, hence, suggest that the costs and benefits of the proposed preferred LT2 option may be lower than that projected using the ICR in the EA.

4. Although a single study and preliminary, further evaluation of those data reveal that "detection of infectious oocyst in filtered water was not related to the historical levels of raw water *Cryptosporidium*" (Aboytes and LeChevallier, 2003). If future research confirms this finding, then this observation suggests that utilities with relatively poor source waters, and who will need to install the most additional equipment and/or take the most other steps from amongst the toolbox options, may incur costs with minimal additional benefit to be derived.

The conclusions reached by EPA using Bayesian modeling are highly dependent on the assumptions made, many of which are not "prior knowledge." These assumptions and opinions should be carefully documented and justified by EPA, and sensitivity analyses should be run and documented in ranges agreed to by a wide array of experts. The effects on the rule and on the benefits of these assumptions should be ascertained, communicated, and considered.

Also, the ultimate objective is to better understand finished water quality. Thus, greater focus on monitoring finished water (rather than analyzing source water and then projecting finished water quality before and after projected treatment) seems to be a more suitable path for future investigation. Limited finished water data suggest that oocyst levels may be more uniform across facilities, regardless of source water, than EPA projects in the EA.

4. Exposure Assessment

Given projected levels of *Cryptosporidium* oocysts in finished water, the next step in the analysis is to determine the level of exposure faced by populations served by rule-impacted systems. This step entails multiplying estimated finished water concentration levels (oocysts/100 L, at levels projected as described in the previous chapter) by the number of liters of finished CWS water a person is expected to drink (L/day). This results in an estimated level of (or distribution of) oocysts/day of exposure across the CWS-impacted population. This chapter focuses on the issue of how many liters of CWS water a person is likely to consume.

4.1 Use of Water Consumption Data from USDA

To estimate the amount of water ingested per day, EPA uses data from the U.S. Department of Agriculture's (USDA's) *1994 — 1996 Continuing Survey of Food Intakes by Individuals* (CSFII). According to the EA, USDA provides two distributions of water ingestion levels:

- "Distribution 1" is based on total water ingestion from all sources, and has a median (50th percentile) level of 1.05 L/day and a 90th percentile of 2.35 L/day. EPA uses the mean from this distribution, 1.24 L/day, as the basis for its national-level risk estimates.

- "Distribution 2" is based on water ingestion from CWS sources, and has a median (50th percentile) level of 0.71 L/day and a 90th percentile of 2.02 L/day. The mean from this distribution is 0.93 L/day (or 75% of the mean from Distribution 1 that is used by EPA as the basis for its national-level risk estimates).

There are several issues associated with EPA's use of the Distribution 1 mean for its national risk assessment, as outlined below. Overall, it seems more accurate and reasonable to use the mean or median value from Distribution 2 instead. This would result in risk estimates and benefits that would be 75% (0.93/1.05) or 57% (0.71/1.05) of the current EPA estimates, respectively, all else equal.

First, Distribution 2 reflects use of CWS waters by their customers. This is a more relevant use level for the regulation that is targeted on water utilities than "water from all sources," which includes bottled water, household wells, household rain cisterns, and household or public springs, in addition to CWS. Water from CWS represent approximately 75% of total daily water intake (U.S. EPA, 2000). Hence, Distribution 2 is by its definition a more suitable fit for CWS-relevant tap water exposure than is Distribution 1.

Second, there are several reasons to believe that even the results from Distribution 2 may overstate daily ingestion levels for *Cryptosporidium*-relevant waters. These reasons include:

- The USDA water consumption estimates that EPA cites include *direct* and *indirect* water consumption. Indirect water is used for final food preparation in the home, and includes water used for coffee and tea, beverages for which the water is boiled or steamed before ingestion (U.S. EPA, 2000). Hot coffee and tea preparation therefore should inactivate any *Cryptosporidium* oocysts that may be present. Therefore, a portion of the Distribution 2 ingested waters would not carry any risk of *Cryptosporidium* infection and illness, even if viable and infectious oocyst were present in the tap water. Mean per capita direct water use of community system water is 0.50 L/day (U.S. EPA, 2000), which is 54% of the total water use mean of 0.93 L/day. Direct use consumption comprises 41% of total CWS water intake at the median, and 63% at the 90th percentile (U.S. EPA, 2000). Thus risk estimates based on total water consumption could be overestimating actual risk substantially, assuming most water consumed indirectly has been boiled. Using the direct water consumption estimates, rather than total consumption, would more accurately reflect potential exposure to potentially infectious *Cryptosporidium*.

- When these data were collected, approximately 13% of daily water intake was from bottled water (U.S. EPA, 2000). In the 7 to 9 years since the USDA data were collected, there has been a considerable and well documented growth in the use of bottled water and in-home filtration devices [i.e., use of tap water alternatives has grown at a 10% annual rate, according to the Water Quality Association (WQA, 1999)]. Therefore, tap water ingestion for many households has declined considerably since the USDA data were collected. A recent EPA-sponsored Gallup poll (U.S. EPA, 2003c) and AwwaRF-sponsored research by Raucher et al. (forthcoming) indicate that as of 2002, 75% of Americans drink bottled water, 14% to 20% drink ONLY bottled water, 37% use in-home filtration devices, and only 49% to 56% drink exclusively tap water in their homes.

- The ingestion levels for both USDA distributions as presented by EPA in the EA (U.S. EPA, 2003a, p. 5-23) are higher than applied or derived previously. For example, for the proposed radon rule, the National Research Council (1999) estimated ingestion risks based on 0.6 L/day. In the arsenic rule, EPA applied a mean ingestion of 1.0 L/day (U.S. EPA, 1999). Tap water intake as estimated and reported by Roseberry and Burmaster (1992) is similar to the older EPA (1999) levels, and is considerably lower than either distribution that EPA now reports and applies (i.e., median of 0.6 L/day, mean of 0.7 L/day, and 90th percentile of 1.4 L/day). This may reflect differences in accounting for direct versus total (direct plus indirect) use of tap water.

Based on the above, it seems very likely that even the use of estimates from Distribution 2 will overstate (perhaps considerably) the level of tap water ingestion that is relevant and applicable to a *Cryptosporidium* risk assessment.

4.2 Sensitivity Analyses

It seems likely that EPA is using an overstated level of relevant tap water ingestion as the basis of its national *Cryptosporidium* risk assessment. We believe EPA is using results from the wrong USDA-based distribution, and that even the use of Distribution 2 is likely to overstate (perhaps considerably) the amount of relevant exposure that is occurring in U.S. households today. This section describes several adjustments EPA should make to improve its analysis.

First, the households (14% to 20%) that no longer drink tap water should be removed from the risk assessment and benefits analysis entirely. This would reduce the national risk and benefits estimates by the same amount (given the near-linearity of the risk and benefits functions applied). In other words, this indicates national results offered by EPA are overstated by 16% to 25% (i.e., 1.0 - 1.0/0.8 = 25%), due to just exclusive bottled water drinkers alone. And, because use of bottled water is growing rapidly in the United States, even higher percentage adjustments may be suitable to reflect the compliance period of 2013 and beyond.

Second, the amount of relevant water exposure in homes that do rely on tap water (at least in part) should be scaled back to reflect the nonrisk-bearing portions of current water ingestion patterns. These include boiled water use (e.g., hot coffee and tea), and the in-home partial use of bottled water or applicable in-home treatment devices. These adjustments might result in typical relevant in-home tap water exposures of 0.5 L/day to 0.7 L/day, or perhaps lower. This lower relevant ingestion rate would apply in the 80% to 84% of homes that remain in the analysis once exclusive bottled water drinkers are properly netted out.

While we have not developed our own empirical estimates within the comment period, the data and logic provided above suggest a more reasonable projection of what exposure level is most applicable. At a minimum, a sensitivity analysis is warranted. For example, using a 0.6 L/day measure would imply that the risk estimates EPA derives based on the Distribution 1 mean overstate individual risks by a factor of nearly 2 or more (i.e., 1.24/0.6 = 2.1).

If one nets out the exclusive bottled water drinkers, and also reduces applicable daily intake for the remaining homes (to half the level EPA uses, as suggested above), then the overall result would be benefits at 40% to 43% of the levels currently presented in the EPA This implies that the EPA national estimates are quite possibly overstated by a factor of 2.3 to 2.5 (i.e., 1.0/0.4 = 2.5), simply on the basis of how much relevant tap water ingestion is assumed.

4.3 Averted Exposures by Sensitive Subpopulations

As described in Chapter 5, the largest component of the benefits estimates is derived from reduced risk of premature fatality among individuals with AIDS and others among the immuno-compromised community.[1] Premature fatality amongst AIDS patients accounts for over 67% of the EPA's estimated total monetized benefits for the proposed rule.[2] However, these estimates are premised on the consumption of tap water — without any additional precautions — by these members of the sensitive subpopulation.

Given the knowledge gained by medical providers, other caregivers, and the sensitive populations themselves since the 1993 Milwaukee outbreak, it is possible that a large percentage of the at-risk population are using boiled water, in-home UV- or RO-treated water, or bottled water, or are taking other precautions against *Cryptosporidium* exposures associated with tap water. While direct evidence of adoption of such averting behaviors has not been documented and such actions are not directly recommended for all AIDS patients, cautions about the risks of *Cryptosporidium* infections and steps that can be taken to reduce infection risks are a feature of information provided by both private AIDS advocacy/education groups (Project Inform, 2003) and the federal government (Centers for Disease Control and Prevention, 2003). Accordingly, the risk posed to these populations is potentially overstated by some unknown degree by the Agency in its analysis.

4.4 Conclusions

A fundamental component of the risk assessment and benefits analysis is the amount of tap water ingested by the public. The discussion above examines several concerns with the assumptions employed by EPA in this aspect of its EA. For example:

- Accounting for exclusive use of bottled water by an increasing number of households would reduce the subsequent risk analysis and benefits results by approximately 20%.

---

1. Sensitive subpopulations include the young, elderly with other underlying illnesses, malnourished, disease impaired, and those with compromised immune systems.

2. Using benefits estimates in the EA, Exhibit 5.24, AIDS victims account for 85% of fatal risk, and fatal risk reductions account for over 79% of total benefits. Combined, this indicates that 67% (85% of 79%) of EPA's projected LT2 benefits are associated with the Agency's projected fatality rate among AIDS patients. Because such a large proportion of EPA's estimated benefits of the proposed LT2 rule are associated with fatality risks among AIDS patients, it seems that much greater public health "bang per buck" could be derived by focusing *Cryptosporidium* risk control measures on AIDS victims and others who are immunocompromised, rather than setting national regulations.

- Adjusting exposures to reflect direct ingestion in the home might reduce the risk and benefits results by another 50%.

- Reflecting averting behavior by AIDS patients and other sensitive populations would directly and proportionally affect the largest component of the monetized benefits estimates.

While the drinking water intake values used in an analysis are not often the focal point of much scrutiny, the values used do have a very sizable impact on the ultimate risk and benefit findings. EPA should reconsider how it has approached this aspect of the EA, and provide improved and more expansive analysis and documentation. The EA needs to be corrected to reflect the issues and data noted here.

5. Dose-Response Assessment: Quantified Health Effects

For its economic analysis of the proposed LT2 Enhanced Surface Water Treatment Rule, EPA develops a risk assessment based on a multistep dose-response model linking expected reductions in the concentration of *Cryptosporidium* oocysts to expected reductions in the annual incidence of adverse health outcomes. The general steps within this risk assessment are presented in the highlighted portion of Exhibit 5.1, which summarizes the general benefits analysis process used for the LT2 rule. In the sections below, we review and critique of each of these steps within the dose-response analysis, in sequence. Finally, we draw overall conclusions and inferences about the most critical issues with the analysis, and about the potential magnitude and direction of any errors that may be embodied in the EPA analysis.

5.1 From Exposure to Infection: Infectivity Assessment

The dose-response sequence begins with an assessment of infectivity. This estimates the probability that a person will be infected by cryptosporidiosis, given the presence of an oocyst in the water they ingest. This in itself entails a multistep process:

1. What are the numbers of all *Cryptosporidium* oocysts — i.e., the exposure — that a person is likely to ingest in a given day? This exposure variable is derived from the steps outlined in the previous two chapters.

2. What is the duration of exposure (n)? EPA assumes this is 350 days per year.

3. What is the probability that an oocyst present in ingested water is infective? The infectious fraction, multiplied by the exposure level (from step 1), provides an estimated expected dose (d) of infectious oocysts ingested per day.

4. What is the probability (r) that an ingested infectious oocyst will generate an infection in a single host human? This infection index is a measure of the inherent infectivity of multiple strains of *Cryptosporidium* organisms in the water being ingested, which EPA refers to as oocyst survival rate (i.e., it is intended to reflect the probability that an infectious ingested oocyst will survive long enough to initiate an infection).

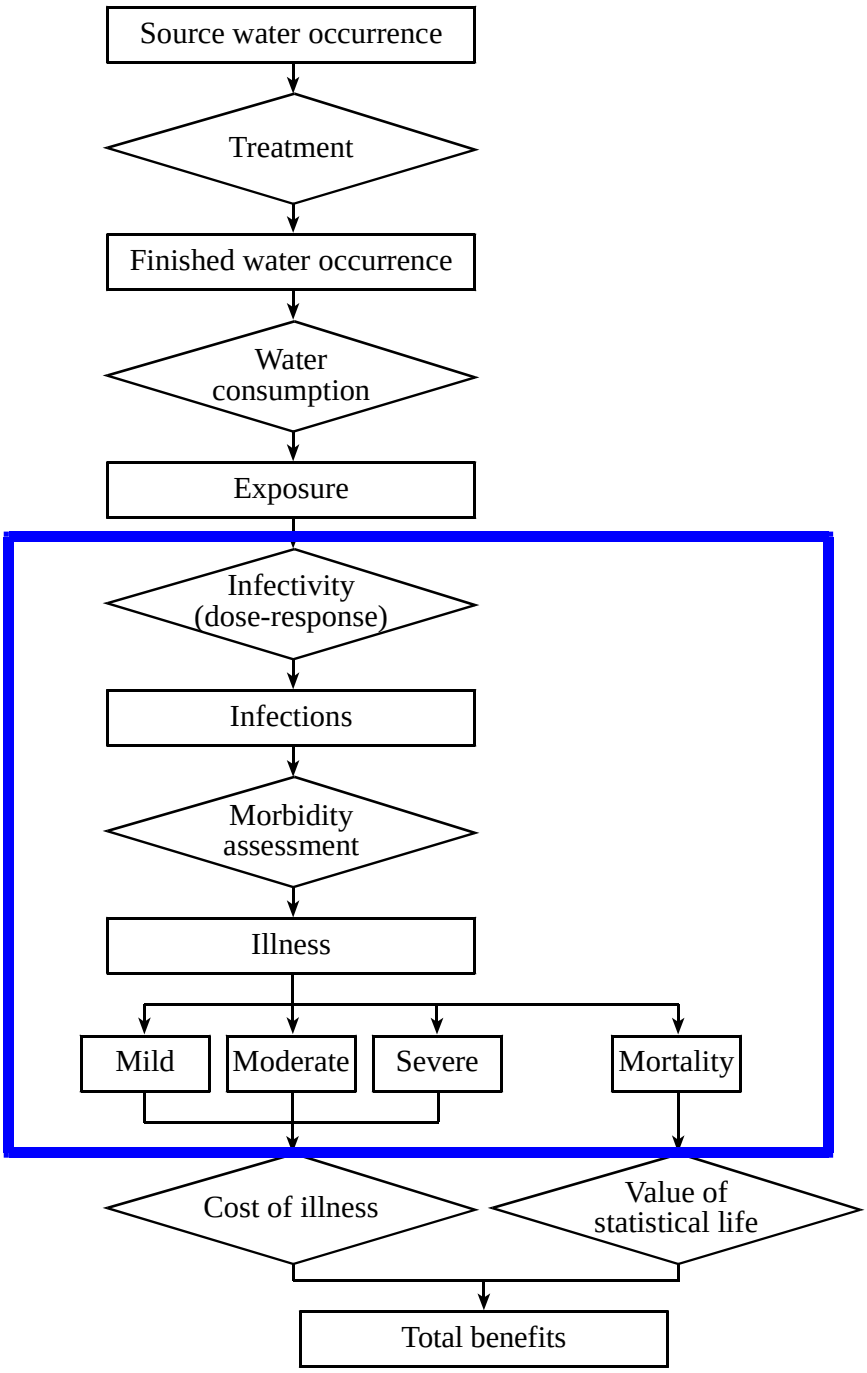


Exhibit 5.1. Risk assessment component (highlighted) of EPA's benefits analysis process for proposed LT2 ESWTR.

Based on the estimates or assumptions applied in each of these steps, annual (or daily) infection risk can be estimated (i.e., the probability of getting infected in any given year of exposure). To do this, EPA uses a simple exponential dose-response model based on Haas et al. (1996). Our review focuses on steps 3 and 4 (since step 1 was addressed previously, and the assumption applied by EPA for step 2 is reasonable).

5.1.1 Infectious oocyst fraction

EPA bases the fraction of oocysts in finished waters that are expected to be infectious on research by LeChevallier et al. (2003) in which oocysts were detected in 60 of 593 samples (using Method 1623) and infective oocysts were identified in 22 of 60 samples (by the CC-PCR method). These results suggest that roughly 37% of *Cryptosporidium* oocysts detected by Method 1623 are viable and are infectious.

To reflect uncertainty in the precision of this implied probability of detected oocysts from the ICRSS databases being infectious, EPA applies a range of values from 30% to 50%, and applies a triangular distribution to this range (so that the mean, median, and mode of random draws from the distribution will approximate 40%). This seems to be a practical and reasonable approach if the 37% finding from LeChevallier et al. is robust. Without additional studies, it is impossible to know if 37% is suitable or representative, but it is the only apparently credible piece of data with which to proceed at this time.

For the ICR data, EPA uses a mode of 20%, with bounds of 15% and 25%. This lower infectious fraction applied to the ICR data is intended to reflect differences in findings due to the lower reliability of the oocyst counts developed using the ICR method (versus the ICRSS findings that were derived using Method 1623). Such an adjustment by EPA seems appropriate in spirit, but we cannot evaluate whether the specific empirical adjustment is suitable or not.

5.1.2 Inherent infectivity — oocyst survival probability (*r*)

Given ingestion of a viable and infectious oocyst, the next step of the dose-response assessment entails estimating the probability that the ingested oocyst will initiate an infection. This is characterized as the variable *r* in EPA's model, and it is the probability that a single ingested infectious oocyst will survive long enough in its human host to cause an infection in the host.

As a probability, *r* can theoretically range from 0 to 1. An *r* value close to 1 is a highly infectious isolate (and an *r* is near zero for strains not very infectious). The value of *r* is thus dependent on what strain of *Cryptosporidium* is ingested, and very limited data are available with which to estimate values for this critical variable. Thus, a core question for this review of the EPA EA is:

"Does EPA develop and use realistic and defensible *r* values?" This is a challenging question to answer, given the sparse discussion and back-up documentation provided by EPA.

There are several issues of general concern with developing *r* values, given the limits of the available data. There also are several additional concerns with how EPA appears to have interpreted these scant data. These concerns are outlined below.

Data are limited to a few strains, and we do not know what strains may be in drinking water

Only three or four strains (the TAMU, IOWA, and UCP strains, plus perhaps one other) have some clinical trial data from which *r* values might be estimated. It is unknown how many strains that are potentially infectious in humans may be present in water supplies, or which of the strains are most likely to be present, or what mix(es) of strains may be present in what proportions. Thus, when oocysts are present (or projected to be present) in finished waters, we typically have no idea if or how the strain(s) present may or may not relate to the three or four strains for which some human health risk data exist. Nor do we know how infectious the other strains might be relative to the three or four strains. Thus, we have no strong basis for projecting real world risks based on limited health risk data available for only three or four strains.

The data available on the three strains used in the analysis are extremely limited

Even for the three strains that EPA uses (TAMU, IOWA, and UCP), our knowledge of the infectivity of these strains is very limited. The available data are based on observations from a small number clinical human experiments. These clinical experiments may be well executed and adhere to good scientific practices. However, each study used very small sample sizes and a limited range of dose levels (generally much higher than applicable for the LT2 baseline and post-rule compliance scenarios).

EPA summarizes the clinical trial findings for each strain in Appendix N of the EA (U.S. EPA, 2003a), and the data are also described in Messner et al. (2001). It is evident that:

1. *Sample sizes were extremely small*. The IOWA strain had 29 subjects tested spread over eight very different dose levels, the TAMU study had 14 subjects spread over four dose ranges, and the UCP strain study involved 17 subjects spread over four dose levels.

2. *Dose levels were extremely high*. The lowest dose level tested was 10 oocysts (TAMU, administered to only three subjects), and the next lowest level was 30 oocysts (five subjects with IOWA, three with TAMU). All other subjects (52 of the 60) were dosed at levels between 300 and 1 million oocysts (19 subjects were dosed at levels of 1,000 or more oocysts). These levels are extremely high compared to human exposures in typical finished waters, which may have infectious oocysts a levels of only 0.00044/L (Aboytes

et al., 2000). If a person directly consumes an average of 0.6 L/day of finished water, it would take over 10 years to expect exposure to one infectious oocyst (and 104 years to have an expected ingestion of 10 infectious oocysts).

3. *Infectivity appears to be positively related to dose*. While the sample sizes are small, the data for each of the three strains show an unambiguous positive association between infection rates and dose levels. Since the dose levels are so high in the studies, one has to question whether the results provide reliable indications of risk of infection when extrapolated to the much lower doses expected in finished waters (where the odds of ingesting an infectious oocyst may be 4.4 in 10,000 per liter, or lower according to the EA).

4. *The fraction of infectious oocysts in the clinical trials may be higher than in finished water*. The clinical trials administered oocysts of a single known infectious strain (either TAMU, UCP, or IOWA) to the subjects. In finished waters through the United States, we do not know what strain(s) may be present or the infectious fractions.

5. *The use of "presumed infections" may significantly overstate the estimated risk.* In Messner et al. (2001), the clinical study data are used to estimate *r* values, using an expanded definition of infection. Specifically, the estimates for infection are based on both demonstrated infections (e.g., the subject is shedding oocysts and shows symptoms) and "presumed infections" (in which the subjects show symptoms but no oocysts are observed in stool samples). For some doses of the clinical tests, these presumed infections make up 50% or more of the total number of infections that Messner et al. (2001) use in their analysis to estimate *r*. Because some of the "presumed infections" may in fact reflect apparent symptoms due to factors other than *Cryptosporidium*, the resulting *r* estimates may be overstated. Given the extremely small sample sizes, any presumed infection that is actually not a case of cryptosporidiosis would have a large impact in the resulting risk estimate.

EPA's basis for "*r*" values used in the EA is unclear and unsubstantiated

There are complex issues regarding how to estimate *r* values from the available data, and which *r* values to use in the analysis. Despite the limitation with the underlying clinical data, *r* values for the three studied strains and an unknown strain have been estimated by EPA researchers in a peer-reviewed publication by Messner et al. (2001). They report *r* values ranging from a low of 0.038% (for UCP) to 0.53% (IOWA) to a high of 4.8% (TAMU). Thus the average across the three strains is 1.8%. They also develop an *r* value of 2.8% for an unknown strain, and this 2.8% has been applied in subsequent studies (i.e., Aboytes and LeChevallier, 2003).

Instead of using these published findings, EPA applies a complex set of modeling techniques to generate estimated probability distributions for r, using two varieties of logit models and running each model using data for either all three strains or for only the two more infectious strains (i.e., dropping UCP from the analysis). The procedures EPA applies are described obtusely in Appendix N of the EA (U.S. EPA, 2003a), and appear to generate distributions of expected *r* values. EPA's means for expected *r* values are estimated to be 7.3% to 8.9% using the logit normal model (with and without UCP in the analysis, respectively), and 9.0% to 10.5% using the logit t model (with and without UCP in the analysis, respectively) (bottom of page N-9). The simple average of these four results is over 8.9%.

It is noteworthy that estimated mean expected values for *r* that EPA derives and uses in the EA are considerably higher than the *r* values reported in the published Messner et al. (2001) article. The average of the EA-reported mean expected *r* values is over 8.93%, whereas the average known strain estimate from Messner et al. is 1.80%. This reflects a difference of a factor of 5 (8.93/1.80), even though both sets of estimates are derived from the same underlying data. The average of the EA results is also above the upper 90th confidence limits for all of the meta-analysis risk factors estimated in Messner et al. (2001).

Given the magnitude of the difference between previously published findings and the results applied in the EA by EPA, there is a need for much greater explanation and documentation by the Agency to describe its rationale and justify its revised findings. It is neither apparent from the EA what EPA's rationale is for the new approach nor very clear what analyses the Agency actually performed, how the results it derived are used, what the range of outcomes are, and how they can or should be interpreted relative to prior published peer-reviewed results. Appendix N in the EA does provide some highly technical information and results, but these are obtuse and lacking in transparency or interpretation. Given the importance of the *r* estimates in the calculation of risks and benefits for the LT2 EA, EPA must provide much more complete and transparent discussion of its approaches, results, and applications.

From EPA's Appendix N, it appears that EPA has applied a complex modeling approach called Markov Chain Monte Carlo (MCMC) in an attempt to overcome perceived and actual deficiencies in the available data. These methods assume that there is information that is understood and from which additional insights can be gained. While these MCMC methods have been extensively applied in the literature, they are computer intensive, fairly complex, and most effective in situations where components of the process are well understood. However, this is not a case in which the processes are well understood. In addition, Appendix N is very difficult to follow, its assumptions are not well documented, and the graphics are also not always clear (including many graphs that do not have labeled axises, e.g., Exhibits N.5 and N.11 through N.16).

The differences between Messner et al.'s results (1.8% to 2.8%) and those in Appendix N (7.3% to 10.5%) can be attributed to EPA's response to the Science Advisory Board (SAB) review. While these SAB suggestions are presented, they are not explained, and the sensitivity to the models is rather dramatic given that changes in the mean results for *r* are in the range of a factor of 4 or 5. This is of concern especially given that the underlying models used [log normal and log t (3df)] are just approximations and there is not solid evidence presented other than the SAB recommendation that these models may be the correct underlying models. Yet the small changes suggested by the SAB resulted in significant changes in the mean values of *r*, which are critical to the general benefits analysis. This is a kind of a sensitivity analysis, and the results suggest that changing the distributional assumptions can dramatically alter the results of the modeling. By the very nature of this process, the SAB and EPA have highlighted that the assumptions made are not perfectly understood. Partially valid arguments can be made for log normal distributions, Student t distributions, and blends of these two. Yet when the different distributions are applied, significantly different results are observed. No one can say emphatically which distribution is correct in the modeling. In fact, it is likely that the empirically true distribution is not one of the standard distributions available to choose from for the modeling effort. Since the outcome appears to be very sensitive to the underlying distributional assumptions, the results of the model should not be considered a good estimation of the dose-response parameter. It would be prudent for the calculations to be done at values of *r* ranging from 1.5% to 11% and to determine the differences based on these differences.

Appendix N needs to be rewritten with much clearer explanations of what is being done, the assumptions being made, how it compares to Messner et al.'s peer-reviewed published paper (2001). The rewrite should also make evident the effects of the assumptions made.

One other fact gleaned from Appendix N that is noteworthy is that EPA recognizes from its review of the clinical studies that infectivity is positively related to dose (reflecting the number of infectious oocysts ingested, as noted on page N-3 of the EA). However, as noted later in this chapter, EPA elsewhere assumes that morbidity is independent of dose (e.g., in its use of the Milwaukee outbreak data to characterize the number and severity of illnesses expected at the much lower endemic exposure rates associated with the LT2 rulemaking). This potential inconsistency with respect to the impact of dose requires greater elaboration and sensitivity analysis by EPA.

5.2 From Infectivity to Illness: Morbidity Assessment

A person infected with cryptosporidiosis does not necessarily become ill — many infections are asymptomatic. Therefore, the probability that a person will become ill (have symptoms) if they are infected must be estimated.

5.2.1 Baseline levels of illness

As a starting point of the analysis, it is useful to assess how many cases of cryptosporidiosis are projected for the regulatory baseline (pre-LT2), and then assess how many of these cases EPA expects will be averted by implementation of the rule. For the baseline, we examined EPA's Regulatory Impact Analysis (RIA) for the Interim Enhanced Surface Water Treatment Rule (IESWTR), which established the baseline for LT2. According to the RIA (U.S. EPA, 1998), the endemic incidence of cryptosporidiosis is estimated to be between 208,500 and 643,000 cases per year, pre-IESWTR, for the relevant impacted systems (the range reflects alternative EPA assumptions about removal rates at baseline treatment).

This baseline is reasonably consistent with the results implied by Mead et al. (1999), as discussed in the EPA EA (U.S. EPA, 2003a, p. 5-6). Mead et al.'s analysis suggests a pre-IESWTR baseline level of 270,000 cases of endemic cryptosporidiosis annually caused by contaminated water. This is based on Mead et al. (1999) using 15 million doctor visits for diarrhea, assigning 2% to cryptosporidiosis, netting out 10% for food-borne cases, and leaving the balance to water and person-to-person contact. If these assumptions are reasonable, then this might be a low estimate given typical under-reporting (i.e., many people with diarrhea do not seek medical assistance). Hence, the range noted in the IESWTR as a baseline seems reasonable.

The IESWTR RIA predicts that the enhanced filtration implemented under the IESWTR will reduce endemic cases by between 149,000 and 432,000 cases per year (U.S. EPA, 1998). This leaves a remaining level of endemic cases of cryptosporidiosis of *between 59,500 and 111,000* per year (e.g., 643,000 minus 432,000 = 111,000) after implementation of the IESWTR. This is the relevant baseline level of estimated cases from which the LT2 rule should begin.

5.2.2 Comparing baseline cases to EPA estimates for cases avoided by the proposed rule

Given this estimated baseline for LT2, it is interesting to note that EPA estimates that because of LT2, the mean number of annual endemic cases of cryptosporidiosis avoided will be *256,000 to 1.02 million* illnesses (and 37 to 141 avoided premature fatalities).[1]

Comparing the estimated illnesses avoided to the estimated pre-rule baseline reveals an obvious problem: the EPA benefits analysis for the LT2 rule is driven by a number of cases avoided that is between 4.3 and 9.2 times higher than the Agency's own estimate of the baseline. Granting that both the baseline LT2 estimate and the estimated reduction due to the proposed rule are subject to considerable uncertainty, the EPA analysis of cases avoided is still problematic. The

1. At the 95th percentile, EPA predicts over 2.3 million cases avoided annually (Exhibit 8.3), or more than twenty times the baseline number of annual cases.

Agency should provide a more complete explanation of both the suitable baseline and the cases it expects to avoid owing to the rule. As it stands, the analysis lacks overall credibility.

5.2.3 Morbidity assessment

To move from its estimate of infections per year, EPA needs to predict how many infected persons actually become ill (i.e., exhibit symptoms). EPA uses triangular distribution to model the morbidity rate, using a mode of 50%, and lower and upper bounds of 30% and 70%, respectively. This is based on human ingestion trials in which, according to EPA, morbidity rates were generally in the range of 39% to 58%, depending on strains used. This suggests that some midpoint between these values — such as the 40% EPA selects — is a practical and logical morbidity rate around which to anchor the uncertainty analysis. However, the 40% level may be too high, given the nature of the underlying data upon which it is based.

As noted above (Section 5.1.2), various critical limitations are associated with the clinical trials, including the small number of subjects and the very high level dosing with known infectious oocysts. Since rates of both infection and illness are believed to be positively related to higher doses — as acknowledged by EPA (EA, p. 5-11) — and since the illness rates reported in the literature pertain to trials using extremely high doses (well above the levels anticipated in endemic exposures in finished U.S. waters post-IESWTR compliance), it follows that the observed numbers of illnesses are higher than would be anticipated at lower, endemic dose levels.

In contrast, EPA argues that the clinical trial studies may underestimate morbidity because the subjects were all healthy and because in some instances only diarrhea was used to indicate illness (possibly ignoring other symptoms). These points may be relevant; however, it seems that the issue of dose rates may more than counterbalance these factors and thus imply some level of underestimation. In addition, EPA's use of the "presumed infection" concept to estimate risk may wrongly attribute some symptoms from the clinical trial subjects to *Cryptosporidium*, thereby overstating the risk. In general, EPA's discussion in the EA seems a bit one-sided. The Agency should re-evaluate the morbidity rate assessment and present a more balanced perspective.[2]

Of particular concern in this stage of the analysis is EPA's assumption that the "morbidity rate is independent of dose . . . the results of this analysis would not be affected by using increased morbidity rates with significantly higher doses" (EA, p. 5-11). Is this indeed a reasonable

2. EPA also needs to consider if the expanded definition of "infected" from Messner et al. (2001), and used by the Agency in Appendix N, affects the application of these data to the morbidity assessment. The inclusion of "presumed infected" may imply that a lesser fraction of "infected" become ill.

assumption? Might an alternative approach be equally or more plausible? If so, what effect would this have on EPA's results in the EA?

Given that infective dose levels in the clinical trails cited in Messner et al. (2001) are at levels between 23,000 and 2.3 billion times greater than observed in finished waters (Aboytes et al., 2000; Aboytes and LeChevallier, 2003), and given that some studies do indicate higher morbidity rates at higher doses (DuPont et al., 1995), it would seem prudent for EPA to explore alternative assumptions to dose independence. These should be included as sensitivity analyses.

5.3 From Illness to Severity

While the dose-response framework incorporated in EPA's risk assessment of the LT2 is well suited for quantifying the expected health impacts of the rule, EPA's execution of the analysis raises a number of concerns and questions. Specifically, there is concern with respect to the process used to account for risks to sensitive population subgroups. This concern includes the general process used to account for adverse health outcomes experienced by sensitive population subgroups and how risks faced by recognized sensitive population subgroups are presented and calculated. Cumulatively, these issues raise questions with regard to the number and distribution, by severity, of adverse health outcomes estimated to be a result of the LT2. These concerns are addressed in this section.

5.3.1 Treatment of sensitive population subgroups in the LT2 risk analysis

In its summary of the risk assessment guidelines, EPA notes that "when the risks posed are not the same for all persons, that variability should be described." Further, the summary of guidelines notes that ideally these risks will be addressed through "the use of scientific data (or reasonable assumptions if data are not available) to produce estimates of the nature, extent, severity, and degree of risk" (U.S. EPA, 2003a; p. 5-4 for both quotes).

For the LT2, EPA initially identifies several population subgroups that are expected or known to face elevated risk of morbidity and mortality following a microbial pathogen infection. These population subgroups include individuals with AIDS, the very young, elderly with other illnesses, and the disease impaired (U.S. EPA, 2003a, p. 5-6). The common feature linking these individuals is that their immune systems are weakened relative to those of persons in the remaining portion of the population through some combination of illness, lack of maturity, or decline with age.

Evidence of the sensitivity of individuals in these subgroups to microbial pathogen infections abounds. Specifically, EPA reports a risk of death of roughly 1 in 100,000 among healthy individuals infected by a microbial pathogen (based on 1% of infections expected to result in hospitalization and a risk of death following hospitalization of less than 1 in 1,000 — U.S. EPA, 2003a, p. 5-7). In contrast, EPA reports a range in mortality risks in AIDS patients following infection of *Cryptosporidium* of 52% to 68% based on data from cryptosporidiosis outbreaks in Las Vegas, Nevada, and Milwaukee, Wisconsin, respectively (U.S. EPA, 2003a, p. 5-14). This elevated sensitivity among the immunocompromised is reflected in the baseline mortality parameter estimates in EPA's dose-response model, where 81% or 88% of the mortality risk applied to the general population is attributable to expected deaths of AIDS patients (U.S. EPA, 2003a, p. 5-14).

Ideally, for the LT2 analysis, EPA would have developed a population-based estimate of the avoided health effects attributable to the rule by calculating and then summing the impacts for identified population subgroups facing elevated risks and then for the remainder of the population. However, despite the apparent availability of population and risk data specific to relevant subgroups (e.g., persons living with and without AIDS), EPA did not pursue this option. Instead, EPA either ignores possible risk differences for these subgroups (i.e., for morbidity outcomes) or uses population-based weights to aggregate subgroup-based risk estimates (i.e., mortality risks) to create a general population risk estimate.

In one respect, EPA's approach could be interpreted as being consistent with the previously summarized risk assessment guidelines because sensitive population subgroups are identified and separate outcome-based risks are developed for the subgroup and the rest of the population. However, the health benefit estimates for the LT2 raise questions when considering how specific characteristics of the critical population subgroup (i.e., persons living with AIDS) that could influence the results are lost in the approach ultimately used.

Specifically, EPA makes adjustments to its mortality risk estimates to try to account for the distribution of persons living with AIDS in the United States, but the adjustment effectively assumes a uniform dispersion of persons living with AIDS in specific types of water systems. The clear limitation to this approach is that persons living with AIDS in the United States are not uniformly dispersed but highly concentrated, as shown in the state-level results in Exhibit 5.2.

Exhibit 5.2 shows that over 30% of the U.S. population living with AIDS in 2001 was concentrated in three states and that 50% of this population was concentrated in only five states. Demonstrating that the concentrations are independent of population are the results on the prevalence of persons living with AIDS per million individuals in the state populations. These values range from a low of 72 in North Dakota to a high of 12,595 in Washington, DC. While the data are not developed here, it is believed that the state level data in fact mask an even greater heterogeneity in the distribution of persons with AIDS, which would be revealed if the state

Exhibit 5.2. Estimated U.S. population of persons living with AIDS as of December 2001

Location	People living with AIDS	Percentage of total population of persons living with AIDS	Prevalence of persons with AIDS per 1 million persons in general population[a]
United States	343,429	100%	1,220
New York	56,792	17%	2,993
California	45,428	13%	1,341
Florida	38,742	11%	2,424
Texas	24,936	7%	1,196
New Jersey	15,702	5%	1,866
Pennsylvania	12,680	4%	1,032
Maryland	11,288	3%	2,131
Georgia	11,269	3%	1,377
Illinois	10,717	3%	863
Puerto Rico	9,548	3%	2,507
Massachusetts	7,368	2%	1,160
District of Columbia	7,205	2%	12,595
Virginia	6,443	2%	910
Connecticut	6,123	2%	1,798
Louisiana	5,851	2%	1,309
North Carolina	5,402	2%	671
South Carolina	5,172	2%	1,289
Tennessee	5,021	1%	883
Ohio	4,905	1%	432
Michigan	4,884	1%	491
Missouri	4,548	1%	813
Washington	4,426	1%	751
Arizona	3,612	1%	704
Alabama	3,427	1%	771
Colorado	3,121	1%	726
Indiana	2,944	1%	484
Mississippi	2,341	1%	823
Nevada	2,249	1%	1,125
Oregon	2,218	1%	648

Exhibit 5.2. Estimated U.S. population of persons living with AIDS as of December 2001 (cont.)

Location	People living with AIDS	Percentage of total population of persons living with AIDS	Prevalence of persons with AIDS per 1 million persons in general population[a]
Kentucky	1,873	1%	463
Arkansas	1,781	1%	666
Minnesota	1,737	1%	353
Oklahoma	1,685	0%	488
Wisconsin	1,669	0%	311
Delaware	1,367	0%	1,745
Utah	1,089	0%	488
Hawaii	1,070	0%	883
New Mexico	1,040	0%	572
Kansas	1,038	0%	386
Rhode Island	961	0%	917
Iowa	623	0%	213
West Virginia	538	0%	298
Nebraska	522	0%	305
New Hampshire	507	0%	410
Maine	486	0%	381
Alaska	239	0%	381
Idaho	233	0%	180
Vermont	216	0%	355
Montana	172	0%	191
South Dakota	95	0%	126
Wyoming	80	0%	162
North Dakota[a]	46	0%	72

a. The population density estimates are calculated using state population estimates from the 2000 Census. The results are unlikely to vary much in terms of magnitude or relative ranking if calculated using 2001 population estimates.

Sources: AIDS population estimates (Centers for Disease Control and Prevention, 2001); U.S. state population estimates (U.S. Census Bureau, 2003).

AIDS populations were allocated by county. The data used to calculate the AIDS-based mortality adjustment provide some indications of this because the data indicate that 14% of those living with AIDS in the United States in 1999 resided in New York City alone (U.S. EPA, 2003a, see Exhibits C.1 and C.2).

Ultimately the heterogeneity in the distribution of persons living with AIDS is important for the LT2 risk assessment. It strongly suggests that accurate estimates of avoided health outcomes, and especially avoided fatalities that are currently driven by outcomes attributable to persons with AIDS, require accounting for expected impacts in the specific water systems serving the population with AIDS.

To emphasize this conclusion, consider the example of implementing the LT2 for a hypothetical population served by two water systems, one that serves persons with AIDS and the other everyone else. In the first case, the LT2 reduces microbial pathogens in each system's water by 2.5 logs, providing a benefit of 10 avoided deaths. Using EPA's current mortality risk estimates, roughly eight of these deaths would be avoided in the AIDS population. In this scenario, knowing where the AIDS population is residing is not important because everyone experiences the same initial risk reduction. However, assuming a linear relationship between avoided outcomes and changes in contamination in a population, consider if the system serving the non-AIDS population were to have its contamination reduced by 5 logs while there was no change in the contamination of the system serving the AIDS population. In this example, the simple average of the change in contamination across both systems is still a reduction of 2-3 logs but now only four lives would be expected to be saved (i.e., a doubling of the lives saved from the original scenario estimate for the non-AIDS population and no lives saved for the AIDS population). As this comparison clearly shows, recognizing risk differences within population subgroups and accurately establishing the distribution of members of those groups among affected systems is a matter of considerable importance in accurately estimating the LT2's health benefits.

5.3.2 Quantification of risk in sensitive subpopulations for the LT2 analysis

Section 5.3.1 argues that an accurate estimate of the health benefits of the LT2, both in terms of number of avoided outcomes and the severity of those outcomes, requires developing risk estimates for identified sensitive population subgroups and for the rest of the population. In general terms, the LT2's risk analysis currently estimates avoided health outcomes by first calculating the number of expected microbial pathogen illnesses in a population. From this illness total, the number of expected deaths are then calculated. Nonfatal cases are calculated as the difference between the original morbidity estimate and the number of deaths, and are then allocated across three categories of severity (mild, moderate, and severe) (U.S. EPA, 2003a).

Starting from the assumption that the number of illnesses has already been estimated, this section demonstrates how the health impact quantification could be completed separately for the population with AIDS and the rest of the population, with data already being used or readily available.

Allocating populations to subgroups

Consistent with the approach partially adopted by EPA, we are interested in defining the population with and without AIDS in the water systems that would be affected by the proposed LT2 rule with as much geographic precision as possible. Clearly, AIDS population estimates are available at a level of precision that allows for assignment of the national population to specific categories of water systems based on the adjustment factors to the Milwaukee data that are proposed in the risk assessment using these results (U.S. EPA, 2003a, p. 5-13). In the suggested framework, the AIDS and non-AIDS population, allocated at least by system type, would serve as the baseline populations from which separate microbial infection estimates are developed.

Calculating mortality risks

Mortality results from the 1993 *Cryptosporidium* outbreak in Milwaukee, Wisconsin, as reported by Hoxie et al. (1997), provide the data currently used to start estimating mortality risks following infection by a microbial pathogen in the LT2 ESWTR's risk analysis.

Hoxie et al. report that the Milwaukee outbreak was attributable for roughly 403,000 cases of watery diarrhea among residents in a five county area likely to have been affected by the outbreak. The LT2 risk analysis incorporates this number as the estimated number of infections from the outbreak. In a review of death certificate data, Hoxie et al. then conclude that the outbreak was attributable for 54 deaths in the affected population, and 46 of the 54 individuals also had AIDS (85.2%). EPA's approach in addressing these data is to develop separate estimates of mortality risks based on the total number of infections and whether or not the individuals who died were AIDS patients.

What would be relevant for the proposed subgroup-specific analyses would be estimates of the mortality rate among those infected who had AIDS and those infected who did not have AIDS. While Hoxie et al. do not provide an estimate of the number of AIDS patients initially infected, the data are apparently available because the risk analysis eventually reports that there was an observed fatality rate of 68% in the Milwaukee outbreak among AIDS patients who experienced a *Cryptosporidium* infection (U.S. EPA, 2003a, p. 5-14).

Further, dividing the number of deaths among AIDS patients attributed to those who experienced a *Cryptosporidium* infection by the reported mortality rate provides an estimate of the number of initial infections in AIDS patients. To calculate a mortality rate specific to the non-AIDS

population, the number of deaths among non-AIDS patients (eight) would be divided by the estimated number of infections in this population [402,932 = 403,000 - (46/0.68)].

With separate mortality and population estimates for the population with and without AIDS, it is worth noting that the subsequent adjustment made to the mortality risk estimates in the LT2 to account for the differences in the population with AIDS in Wisconsin relative to these populations in filtered and unfiltered systems nationwide would no longer be required.

Finally, it is worth discussing the adjustment to the AIDS-based mortality rate estimates currently made in the LT2 risk analysis to reflect improved survival of AIDS patients over time (U.S. EPA, 2003a, p. 5-13). The argument presented for this substantial adjustment (i.e., it reduces the AIDS-based mortality rate per 100,000 infected by *Cryptosporidium* by a factor of approximately 25%) is based on the impact of recently developed AIDS medications.[3] It seems that an argument must be presented that the new medications provide some sort of protective benefit against either infection from microbial pathogens or mortality once infected; such justification is not presented.

Allocating nonfatal cases across population subgroups and severity classes

Clearly data currently available and incorporated in the risk analysis could be used to develop separate estimates of the avoided fatal cases of infection by microbial pathogens for a population subgroup of individuals with AIDS and for the rest of the population. Subtracting these fatal case estimates from the initial estimates of illnesses in each group yields the total number of nonfatal cases for each population subgroup.

Currently, the risk analysis for the LT2 makes no allocation of nonfatal cases between population subgroups. Consequently it makes no distinctions by population subgroup in allocating these infections across severity categories using Corso et al. (2003) estimates for the Milwaukee outbreak that 88% of all infections did not require medical attention, 11% were seen as outpatients, and 1% required hospitalization to distribute nonfatal infection estimates across the mild, moderate, and severe outcome categories.

While these estimates are believed to be an accurate representation of the disposition of infections in the Milwaukee outbreak, they would need to be revisited for use in an application where separate nonfatal infection estimates are calculated for population subgroups that either have or do not have AIDS. Specifically, while the distribution is probably appropriate for use with the population that does not have AIDS, it seems questionable for use with the AIDS population given the lack of available treatments and the inability of that population to mount an effective resistance to the infection given their immunocompromised status. Essentially then

3. The overall general survival rate for AIDS patients is adjusted by EPA by a factor of roughly 8.

what would be needed is a separate distribution for nonfatal cases for AIDS patients. Very likely, this distribution would be heavily weighted toward nonfatal infections being in the severe category.

Conclusions

Clearly, given the currently available data, it is not beyond the scope of EPA to develop estimates of the health impacts of the proposed LT2 based on defining population subgroups that have different health risks with respect to microbial pathogens. While doing so would clearly increase the precision of the estimates, it is not currently clear in what direction the current estimates may be biased given the multiple factors that would need to be accounted for, especially linking actual subgroup populations to specific utilities that would be affected by the rule.

Finally, it is worth noting additional items that could and should be addressed qualitatively, if not quantitatively, in the risk analysis. This is regardless of whether the alternative of developing population health impact estimates by estimating totals for mutually exclusive population subgroups and then summing is pursued.

First, EPA's risk assessment currently assumes one distribution for a morbidity factor that estimates the probability of developing an illness if infected by a microbial pathogen (U.S. EPA, 2003a, p. 5-11). A discussion of how this risk is likely to vary between those who are immunocompromised and the rest of the population is warranted given the clear identification of this at-risk group.

Second, the risk assessment does not discuss averting behaviors that individuals could take to lower their risk of exposure to infective microbial pathogens. For most of the general population, it could be presumed that such steps are rarely taken. However, a discussion with respect to the identified sensitive subgroup of immunocompromised individuals seems warranted. Such a discussion could evaluate standard recommendations provided to those with AIDS by support groups or medical professionals on best practices for using tap water and/or recommended preventative measures to reduce exposure to microbial pathogens. While it is not clear that such measures are widely incorporated by members of this population, it is fairly clear that outbreaks such as that for *Cryptosporidium* in Milwaukee in 1993 have heightened awareness of microbial pathogen risks among this population.

5.4 Use of Outbreak Data to Infer Inputs and Outcomes for Endemic Risk

A question of significance not addressed in Section 5.3, but one that plays a critical role in the calculation of expected health benefits for the LT2 ESWTR once the number of infections from exposure has been calculated, is the reliance on data from outbreaks of microbial pathogens to estimate fatal cases and the severity of nonfatal cases.

Ideally, data on the impacts of endemic exposure to microbial pathogens would be used to estimate the impacts of the LT2, but such data do not exist. A reasonable question then is whether the use of data related to outbreaks such as that for *Cryptosporidium* reported by Hoxie et al. (1997) and Corso et al. (2003) for Milwaukee's 1993 outbreak may bias the results and in what direction.

Results based on data from outbreaks used to estimate impacts on endemic illness may be biased upwards if the response rates and the severity of the response increase with the level and duration of exposure. This is a possibility EPA recognizes clearly in discussions on its mortality risk estimates (U.S. EPA, 2003a, p. 5-14). Second, there may be an upward bias in the observed health responses during an outbreak if events that would have received a different medical code ordinarily are coded to imply attribution to the outbreak, in part because of the awareness of medical providers of the event itself. Further, given the general conclusion that most individuals will not seek care for mild microbial pathogen infections, it is very possible that the estimates of severe impacts based on estimated numbers of infections and case distributions observed during an outbreak may be upwardly biased if the true number of infections is much larger (e.g., due to greater awareness and concern by those mildly affected).

Unfortunately, this is a data shortcoming that is currently difficult to address. Because it is unlikely that detailed and reliable information on the effects of endemic exposures will be developed, the only alternative to using current data inputs is to incorporate additional sensitivity analyses designed to evaluate the extent to which assumptions would have to change before critical decisions were altered. Once such values are identified it may be possible to evaluate their plausibility in order to provide additional information for consideration.

5.5 Conclusions

The dose-response portion of the EA is a very complex and highly significant component of the overall analysis. Critical assumptions must be made at several points in the analysis because of pervasive scientific uncertainties. EPA makes several plausible and reasonable assumptions and inferences, but there also are components of the analysis where alternative assumptions or

scenarios seem more plausible or, at a minimum, equally plausible. In such instances, the Agency should be more explicit and balanced about assumptions it is making, and should conduct meaningful sensitivity analyses to reveal the impact of the alternative assumptions on the overall findings.

Of particular concern from our review of the EA are the following:

1. The estimation and use of *r* values for inherent infectivity. The estimation process is not well documented and is poorly presented, the underlying data have significant limitations, and the results that EPA apparently uses are appreciably higher than findings published in the peer-reviewed literature. This requires greater documentation, discussion, and review. The inputs and assumptions that feed the model are also nonstandard (including the use of presumed infections instead of the standard long established strategy of using confirmed infections). The effects of this important assumption on the outcome of the model have not been presented. A kind of sensitivity analysis has been done for the importance of the underlying model and the change from a normal to a t distribution has had the effect of changing the mean value of *r* by a factor of four or five. This demonstrates the importance of the underlying assumptions regarding the distributions used, and requires greater explanation and justification.

2. The morbidity assessment is dose-independent. Given the extremely high doses used in the clinical trials relative to the levels in finished water, and the evidence of dose-dependent morbidity in some studies, it seems prudent to at least conduct reasonable sensitivity analyses on the impact of this key assumption.

3. EPA's projected number of cases avoided is implausible given the baseline level of illness projected by the Agency. In fact, EPA's estimates of avoided cases are between 4 and 10 times higher than its baseline.

4. Morbidity levels, the allocation of cases across severity classifications, and the characterization of the duration and impact of illnesses in each severity class are all based on data from the Milwaukee outbreak of 1993. It is problematic to use data from a massive outbreak (with very high levels and durations of primary and secondary exposure) as a basis for estimating numbers, severity, and duration of illness from endemic exposures to far lower doses. Here again, given the core uncertainties and the importance that these assumptions have on the final benefit outcomes, some alternative scenarios should be developed and assessed using sensitivity analysis.

6. Valuation of Health Impacts

The valuation of the fatal and nonfatal impacts expected to be avoided as a result of the LT2 ESWTR is a critical component of estimating the benefits of the rule because it provides a basis for comparing the benefits of this rule to other potential actions and to the anticipated costs for implementing and complying with the rule. In the valuation process, different data sources and methods are used to monetize estimated nonfatal and fatal infections. In this process EPA follows established precedents and treatments of data in valuing the nonfatal infections using what is called the Traditional cost-of-illness (COI) method, and in its valuation of fatal outcomes using a well established distribution of estimates for the value of a statistical life (VSL estimates). In this regard it is only EPA's development and use of what is called the Enhanced COI method to provide an alternative valuation of the nonfatal infections that merits significant discussion because it is a considerable departure from standard COI practices and does not appear to have been subjected to reliable peer review.

6.1 Use of the Traditional Cost-of-Illness Method to Value Nonfatal Infections

As recognized in the LT2 EA, the preferred measure for valuing any adverse health outcome is a measure of an individual's *ex ante* (ahead of time) willingness to pay (WTP) to avoid the outcome (U.S. EPA, 2003a). WTP measures for valuing an avoided health outcome are an economist's preferred measure because they allow the individual to incorporate information about the expected direct financial costs of the outcome (e.g., medical expenses, lost income, lost household production) along with those features of the outcome that are not directly valued, such as any associated pain and suffering or activity limitations, both short and long term, that may result from the outcome. Finally, because *ex ante* WTP responses are necessarily bounded by an individual's income, they are constrained by the same limits that are faced when making other economic decisions.

Unfortunately, WTP estimates exist for only a limited set of nonfatal adverse health outcomes and this pool does not include outcomes similar enough to those experienced from nonfatal microbial pathogen infections to be suitable for use in a benefits transfer for the economic evaluation of the LT2's expected health benefits. In these situations, one option is to develop and use COI estimates that account for the direct medical expenditures, lost income, and lost household productivity associated with the outcome. Clearly, such COI estimates provide a lower bound valuation estimate with respect to WTP estimates because the method fails to account for the value an individual would place on avoiding any pain and suffering, as well as any activity limitations associated with the condition. Because of the general availability of the

data required to complete COI estimates and their known relationship as a lower bound to WTP estimates, COI valuations are a useful tool for monetizing expected reductions in adverse health outcomes such as those anticipated from the LT2.

The Traditional COI estimates incorporated in the LT2's economic analysis accurately follow the standard COI method by accounting for the types and level of medical services and expenditures associated with nonfatal microbial infections according to their severity level, along with estimating the associated number of productive days that would be lost with each severity level of infection. The Traditional COI results incorporated in the LT2's economic analysis reflect results originally developed in Corso et al. (2003), updated from their 1993 base year values. The Corso et al. values were developed for different severity categories of nonfatal *Cryptosporidium* infections observed in the Milwaukee, Wisconsin, outbreak of 1993, based on a review of medical records and application of assumptions about some required types of care (e.g., number of physician visits). Because the criteria used by Corso et al. to assign nonfatal outcomes to severity levels were incorporated directly in estimating the distribution of nonfatal cases in the LT2's economic analysis, there are no issues with respect to the valuation estimates reflecting a different set of conditions than were used to assign outcomes by severity category. Further, because the costs are developed for infections for one of the key microbial pathogens being targeted by the LT2, they are based on outcomes relevant for consideration in the rule.

The estimates developed by Corso et al. (2003) are noteworthy for the care that has been taken to accurately estimate actual expenditures and the value of productivity losses. For example, the information on hospital and emergency room charges that was initially available was adjusted using a regional cost-to-charge adjustment ratio to reflect the fact that hospital charges are rarely what are paid by the patient because of negotiated discounts with the hospital (e.g., for patients with private or public insurance). In this regard, EPA's Traditional COI valuation estimates can be reliably interpreted as a lower bound estimate on the value of avoided nonfatal infections.

6.2 EPA's Enhanced COI Estimates

The Enhanced COI estimates developed by EPA build on the Traditional COI estimates by expanding on the valuation of the estimates of the productive and leisure time affected by a nonfatal microbial infection (there is no adjustment to the estimated medical expenditures). These differences can be easily summarized in table form and are presented in Exhibit 6.1 (U.S. EPA, 2003a; see Appendix K, Exhibit K.2, p. K-9).

Exhibit 6.1. Differences in the valuation of time in the Traditional and Enhanced COI methods

Time category	Traditional COI valuation	Enhanced COI valuation
Paid work time	Median pre-tax wage plus benefits	Median pre-tax wage plus benefits
Unpaid work time (household production)	One-half median post-tax wage	Median post-tax wage
Leisure time	Not valued	Median post-tax wage
Sleeping time	Not valued	Not valued

As Exhibit 6.1 shows, the fundamental difference with the Enhanced COI method relative to the Traditional COI method is that it both doubles the value of time for lost household production and assigns a value to impacts on leisure time that previously were not monetized. The leisure time values are especially of concern, since they are at odds with standard practice in the economics profession with regard to recreational activity valuation. In recreation demand modeling, it is typical to use only a fraction (e.g., one-third) of the prevailing wage rate to infer the value of time spent traveling to a recreation site (and time spent on-site engaged in the recreational activity generally is not counted as a cost in the demand estimation via the travel cost approach). In contrast, EPA's Enhanced COI approach applies full wage rates to all waking hours. Appendix B provides additional discussion of the problems associated with values EPA assigns to time away from work in the Enhanced COI approach.

The effect of the Enhanced COI method is to increase the average (i.e., severity weighted) value per nonfatal infection to $745 (2000 dollars) compared to $245 for the Traditional COI. In short, the Enhanced COI roughly triples the average value per nonfatal infection from a microbial pathogen.

In placing the Enhanced COI method and results in context, EPA notes that studies that compared WTP estimates for an adverse health outcome with COI estimates developed using essentially the Traditional COI method have produced values ranging from 2 to 79, and that many of the ratios fall in a range of 3 to 6 (U.S. EPA, 2003a, p. 5-42). Implicitly, this provides a justification for the apparently more complete valuation of time incorporated in the Enhanced COI and seems to suggest the results are some sort of reliable proxy for WTP estimates.

It is because of the implicit comparisons that are made with the Enhanced COI results that great care should be taken in their presentation and consideration. The Enhanced COI estimates should never be taken as a proxy for WTP estimates for these outcomes. The only result that can be expressed with confidence about the Enhanced COI results is that they exceed the Traditional COI estimates. Any direct or implied comparisons with what WTP estimates would provide

would be misleading because there is no basis to say whether the Enhanced COI estimates are higher or lower than WTP.

Finally, it is worth noting how some of these issues are addressed in other regulatory analyses prepared by EPA. For example, in its valuation of hospitalizations attributable to concentrations of airborne particulate matter, EPA currently uses a COI approach that considers only hospital charges (not adjusted for the ratio of costs to charges) and time spent in the hospital valued at the median pretax daily wage (Apelberg et al., 2003, p. 4-35). Using unadjusted charges probably overstates the medical costs of hospitalization, but follow-up medical costs are not included so the amount of overstatement is uncertain. This treatment of the value of time gives full wage value to essentially eight hours a day regardless of the person's employment status or the day of the week. On average, about half of these hours are accounted for by paid employment, leaving about four hours per day valued at the wage rate to cover the value of lost household production and leisure. This is substantially less than counting all waking hours at the wage rate as is done in the Enhanced COI approach.

6.3 Placing the Enhanced COI Estimates into Context

One of the key concerns with the EPA's Enhanced COI estimates is that it is difficult to gain perspective on whether the resulting numerical levels "make sense" in terms of what might be considered credible as an indicator of the more suitable WTP measure. This section attempts to develop some perspective on this issue.

First, it is important to recognize that the suitable measure for valuing a reduced risk of a cryptosporidiosis illness is *ex ante* WTP (i.e., what a person would be willing to pay, ahead of time, for an opportunity to reduce a low level risk by some degree when provided with an accurate description of the risks and consequences of the illness). When WTP estimates are not available, a Traditional COI is often used as a convenient but less suitable conceptual and empirical substitute and as a recognized lower bound proxy for value.

EPA's Traditional COI estimate excludes some important value components (e.g., pain and suffering), but COI estimates also are an *ex poste* measure of what costs a realized illness imposes on an afflicted person. The former feature allows analysts to consider that a Traditional COI will provide a lower bound estimate of *ex ante* WTP (i.e., it is the cost of being ill, not the value of reducing the likelihood of illness). But with EPA's aggressive accounting in its Enhanced COI estimate, one loses the ability to assess how the Enhanced COI figure compares to the unknown but more suitable *ex ante* WTP value. With the Enhanced COI, we no longer can say with confidence that incurred costs are less than or equal to the value of reducing the probability of potential future illness — the enhanced approach removes our ability to interpret the COI numbers relative to useful boundaries.

To help evaluate whether the Enhanced COI provides a numeric result that seems plausible and reasonable — vis-a-vis *ex ante* WTP — a simple illustration is developed below using VSL estimates to provide some context. VSL estimates reveal the *ex ante* WTP of individuals to reduce risks of premature fatality. A VSL of roughly $6.3 million saved is often used as a measure of the value of reducing risks of premature fatality, based on a large body of well-reviewed literature in which individuals (e.g., median aged workers) in effect reveal a WTP to reduce a mortality risk typically in the range of 1 in 10,000 per year. What this literature actually tells us is that a typical, median aged person has an *ex ante* WTP of $630 per year, on average, to reduce a 1 in 10,000 per year risk of premature fatality in the coming year. This widely accepted $630 per person of *ex ante* WTP to avoid a 1 in 10,000 annual risk of immediate fatality thus serves as our frame of reference.[1]

According to EPA's analysis of the endemic risk of water-borne cryptosporidiosis, the baseline annual risk is also on the order of 1 in 10,000 (U.S. EPA, 2003a, Chapter 5).[2] Compared to the VSL context, however, the outcome is far less severe — i.e., the risk of getting an illness that, for

1. In deriving the VSL estimates from premature fatality risks in occupations, the wage risk studies are based on annual risk of fatalities that typically range from $1*10^{-4}$ to $2*10^{-4}$. Data from the U.S. Bureau of Labor Statistics and the National Institute of Occupational Safety and Health, as summarized and reported in Viscusi and Aldy (2003), indicate the average annual rates of fatality are about $3.6*10^{5}$ in manufacturing, $1.5 *10^{-4}$ in transportation and utilities, and $2.5*10^{-4}$ in the relatively risky mining sector. Viscusi and Aldy found a mean annual risk of occupational fatality of $2*10^{-4}$ from the numerous risk-wage studies they include in their meta-analysis of VSL.

Viscusi and Aldy (2003) also review more than 60 risk-wage studies, and report the findings of a meta-analysis conducted with 49 suitable studies. Their meta-analysis derives a mean VSL of $6.7 million (2000 U.S. dollars) and a mean occupational risk level across the underlying studies of $2*10^{-4}$. The VSL estimate is based on wage premiums relative to less risky jobs, and assuming the implied risk reductions are about 1 in 10,000 (i.e., that the comparison is on average to jobs with half the fatal risk as the chosen occupation), this suggests an annual WTP premium of $670 per impacted risk-bearing worker.

Several factors could drive these estimates per person annual risk premiums up or down. For example, because over half (25 of 49) of the underlying studies also embedded risk premiums for nonfatal (as well as fatal) injury, the resulting WTP estimates are overstated for the risk of premature fatality alone. In addition, if the risk differential between the risky and less risky occupations in the study differed by more (less) than 0.0001 per year, then the implied annual per person value would increase (decrease).

2. The median annual risk of cryptosporidiosis illness as estimated by EPA is about $1*10^{-4}$ at baseline (pre-LT2). This is based on EPA's Exhibits 5.12 and 5.13, wherein a $1*10^{-4}$ or higher baseline annual risk of illness is anticipated for 46% of the persons served by filtered systems, and "essentially the entire population served by unfiltered systems" (U.S. EPA, 2003a, p. 5-36). The exhibits suggest that over one-third of filtered system customers, and well over 95% of unfiltered systems customers, face annual risks of twice that level (i.e., $2*10^{-4}$ or higher).

99% of the afflicted, would entail a mild and short-lived period of discomfort and diarrhea (rather than immediate death as would apply in the VSL estimates).[3]

To facilitate comparison to the VSL, assume for this illustration that the risk posed by cryptosporidiosis was permanent and that symptoms were suffered through a recurring cycle of typically mild or moderate cases (with perhaps a rare severe episode as well, but no risk of fatality). Thus, a median aged worker struck by the disease under this assumption would face predominantly mild illness (and the associated COI) for the balance of their lives. What numeric result would the EPA Enhanced COI assign to such a case?

- First, the value of time lost for the individual, based on EPA's approach, would be $217.86 per day while they were employed ($18.47 per hour for 16 waking hours). Assuming the median aged person was 38 years old at onset and retirement is age 65, then the value of time lost through their remaining work period would be $2.15 million (27 years times 365 days per year times $217.86 per day).

- Second, add the value of time lost due to illness through the retirement period (age 65 through a typical conditional life expectancy to age 83). Using EPA's estimate of $10.92 per hour for 16 waking hours per day for non-employed individuals, a value of time lost per person of $1.15 million is derived (18 years times 365 days times $174.72 per day).

- Third, the EPA Enhanced COI would add additional costs due to care-giving efforts and medical expenses for each typical nonfatal case. For a typical nonfatal case, if extended throughout a lifetime from age 38 onwards to age 83, these costs would amount to $1.25 million.

Combining these three Enhanced COI elements, the total cost associated with a permanent case of cryptosporidiosis, incurred beginning at age 38, would thus be $4.55 million.

If one were to contend that EPA's Enhanced COI approach provided useful and reasonable approximations of *ex ante* WTP to avoid a lifelong case of cryptosporidiosis such as depicted in this example, then this implies an average *ex ante* WTP of $455 per person exposed to the 1 in 10,000 annual risk of this hypothetical life-long version of the disease outcome.

3. EPA assigns 88% of cases to the "mild" category, in which (based on EPA's use of data from the Milwaukee outbreak) the illness lasts 4.7 days on average and entails a mean loss of 1.3 work days, and 11% of the cases to the "moderate" category, in which EPA believes illness lasts a mean of 9.4 days (3.8 mean days lost from work). The remaining 1% of cases are labeled severe, and entail 34 days of mean illness (a mean loss of 5.4 days of work).

While this illustration uses a somewhat contrived version of cryptosporidiosis (lifelong continuous recurrence of a typical, severity-weighted nonfatal case), it does provide useful context for considering the valuation issue. A lifelong cycle of diarrhea and other mild symptoms would no doubt be unpleasant, and no doubt a typical person would be willing to pay a considerable amount to reduce the risk of such an outcome. However, would they have a WTP that is over 72% of their WTP for a comparable level of risk of immediate fatality (i.e., \$455/\$630 = 72.2%)? This does not seem to be plausible or likely.[4]

Therefore, this numeric illustration reveals that EPA's Enhanced COI approach does not appear to yield plausible estimates of the value of reducing the risk of cryptosporidiosis. The EPA approach, if applied to a hypothetical lifelong case, would result in an estimate of value that is implausibly high relative to better understood risk avoidance values (i.e., VSL).

6.4 Conclusions

EPA's use of COI-based estimates to stand in for the *ex ante* risk reduction WTP for cryptosporidiosis is understandable, given the lack of reliable data from which to infer the preferred, conceptually appropriate measure of value (i.e., *ex ante* WTP). The Agency's Traditional COI is a reasonable estimate for EPA to use in this regard, and these cost estimates may reflect a lower bound as a proxy for value. However, the development and application by EPA of the so-called Enhanced COI approach seems to be lacking in credibility and plausibility. This approach does not adhere to standard practice in the economics profession (e.g., with regard to valuing time spent out of the workplace), and it generates results that do not appear to be reasonable relative to other benchmarks (e.g., VSL) under simple illustrations that can be constructed. Therefore, the Enhanced COI should not be used to evaluate this rulemaking, and should instead be subjected to far greater peer review and revision before EPA applies it to any future matter of public health policy-making.

4. The above Enhanced COI illustration does not account for either real income growth over time (EPA applies a real rate of 2.3% in the EA), nor does it account for discounting future values to reflect the real rate of time preference (EPA uses a 3% real rate of discount to reflect time preference). When the above Enhanced COI estimate for a hypothetical lifelong case of typical severity cryptosporidiosis is adjusted to reflect real income growth and discounting over the applicable timeframe (at 2.3% and 3.0%, respectively), the resulting present value is nearly \$3.1 million per case. This is somewhat less than the undiscounted, zero income growth estimate provided in the main text above, but still disproportionately high relative to a VSL of \$6 million. The same implication arises as above — is it plausible and realistic to assert that the WTP to avoid a case of lifetime cryptosporidiosis (valued at over \$3 million using EPA's Enhanced COI approach) would be half as great as the well studied WTP to avoid a like-sized risk of premature fatality? The answer, we believe, is clearly "no." The Enhanced COI estimates are not likely to be plausible or realistic as proxies for WTP.

7. Compliance Costs and Net Benefits

This chapter provides a cursory review of EPA's approach and results for estimating the costs of compliance with the proposed LT2 rule. The chapter then reviews how EPA compares estimated benefits to costs, with a particular focus on the extent to which an appropriate disaggregated incremental net benefits perspective is followed by the Agency.

7.1 Compliance Costs

National level compliance cost estimates

EPA estimates that the annualized cost of compliance for the preferred regulatory option will be from $73 million to $111 million (using a 3% interest rate to annualize capital costs), with the range dependent on whether the ICRSSL or ICR data were used (U.S. EPA, 2003a, Exhibit ES.6). The costs developed based on the ICR data are higher than the costs based on either ICRSS dataset because of higher allocations of systems to higher bins.

EPA is fairly certain that compliance costs will approximate these estimates, as reflected in its 90% confidence ranges that are only plus and minus 11% of the values reported above. For example, $65 million and $82 million are the reported 5th and 95th percentile cost estimates, respectively, around the central estimate of $73 million derived using the ICRSSL database to estimate occurrence.

We have not conducted an in-depth review of the cost models and assumptions used to develop these estimates, but we find it difficult to believe that EPA is 90% confident that it has accurately forecast the LT2 compliance costs within "11% for any given ICR-related database occurrence profile. There are many complex and uncertain factors underlying the cost analysis, including:

- the extent to which utilities will mix and match elements of the toolbox, and the inherent difficulty of predicting compliance choices to be made by utilities
- unit treatment costs for both capital outlays and operations and maintenance (O&M), for each of the viable compliance options, for each utility size and source water
- site-specific retrofit challenges and associated costs
- the allocation of costs of some utility compliance efforts between the IESWTR and LT2

- the potential for price spikes for UV light disinfection units caused by regulation-driven increases in demand (i.e., the cost ramifications of a possible sellers' market).

Given these and other issues, EPA's stated confidence intervals seem unrealistic. If EPA is going to generate confidence intervals (which it should), then it needs to consider and reveal clearly what levels of realism and precision apply at each stage of the cost analysis, and track how variabilities and uncertainties can compound across stages of the assessment. The cost estimates provided by EPA, despite their presentation of confidence bounds, do not realistically portray the expected level of uncertainty for potential compliance costs.

Finally, within the costing analysis, there seems to be an error in the ozone contactor times EPA reports for 2 log removal. EPA should review and correct this apparent error.

Per household cost estimates and affordability

EPA estimates that the cost per household of LT2 compliance will in general be fairly modest, at $1.07 and $4.61 per household per year, at the mean (Exhibit ES.7). The lower estimate reflects mean costs averaged for all systems (and ICRSSL database outcomes), and the upper estimate pertains to ICR-based costs in systems serving less than 10,000 (U.S. EPA, 2003a). The mean household level cost increase would be between $7.30 and $10.45 per year in "small systems" (less than 10,000 served) needing to add treatment, and fewer than 0.3% of households in this category would be expected to face an increased annual water bill of over $120 (U.S. EPA, 2003a, Exhibit J.4).

A critically important consideration related to interpreting the above EPA results is that the EA designates any CWS of 10,000 or less served as a "small system." The reported per household costs thus reflect a very large range in system sizes (e.g., from 25 served to 10,000 served). This reflects a range that varies by a factor of 400 (10,000/25) in population served within this "small system" size category. Economies of scale can be appreciable over this size range, and the aggregated results depicted by EPA are likely to mask much higher per household cost burdens through the smaller end of the CWS size spectrum.

The EA for LT2 does not appear to provide any information whatsoever about how costs per household vary across the various small system size categories. This lack of transparency and detail should be rectified for the EA that accompanies the final rule. One data point available from the EA for the proposed Stage 2 rule for disinfection byproducts reveals how important this small system disaggregation can be for considering household-level cost impacts. In the Stage 2 affordability analysis, the mean annual per household increase in costs for that proposed rule is nearly 6 times greater in systems serving 500 or fewer people than the mean for households served by systems in the 3,300 to 10,000 served end of the range (and 5% of the households in the under 500 served category will face annual costs projected by EPA to be over 12 times

higher). In addition, for customers in systems of under 100 served, the mean per household costs probably will be considerably higher than predicted by EPA for systems of up to 500 served. EPA does not reveal the costs on a CWS under 100 served basis, but even the limited disaggregation found in the Stage 2 EA reveals how much important information is masked under the inappropriate aggregation of all CWS of 10,000 served or less within the "small system" category.

The key point here is that EPA should provide more finely disaggregated cost results. This is important because there are key equity and affordability implications that are masked (hopefully unintentionally) by EPA under the approach the Agency uses in the LT2 EA to portray its cost and affordability findings. Since the estimated compliance costs are developed using the typical nine system size category scheme EPA has employed in the past, the results are all generated by (and available to) the Agency at that level of disaggregation. Merging the five smaller size categories into one over-broad "small system" category of 10,000 served or less is an extra step made by the Agency, and one that obscures rather than informs public review and Agency decision-making.

7.2 Comparison of Benefits to Costs

It is vital that EPA provide suitable and informative comparisons of benefits to costs. The suitable framework is to reveal incremental benefits, incremental costs, and incremental net benefits (i.e., incremental benefits minus incremental costs) for each relevant regulatory alternative. The increments should start with the suitable regulatory baseline, and move to increasingly stringent (costly) alternatives. Ideally, the preferred regulatory option would be identified where the last incremental net benefits is still positive (i.e., just before incremental net benefits turn negative). To EPA's credit, there is some incremental net benefits information in the LT2 EA. This type of outcome should continue to be a routine and highlighted portion of all EPA EAs in the future (unfortunately, it is lacking in the Stage 2 EA).

Further, it is important that the incremental net benefits be provided not only at the national aggregate level, but also according to informative system size categories (U.S. EPA, 2001). That is, incremental net benefits should be reported for each of the nine CWS size categories EPA uses to build up its cost and benefit estimates. Additional disaggregation is also worth portraying where important distinctions are reflected in costs, benefits, or both. As noted below, other relevant levels of disaggregation include separating results for unfiltered systems from those derived for filtered systems.

Regrettably, EPA fails to provide any meaningful disaggregated benefit-cost comparisons along the lines described above. All benefit and incremental net benefit estimates are shown only at a national aggregate level, with no disaggregation by system size. Cost estimates are at times

disaggregated, but only across very broad (and therefore meaningless) categories based on whether they serve 10,000 or fewer people or over 10,000 people. Overall, the EA is severely disappointing in this regard, and completely at odds with both best practices and recommendations issued in association with past critiques (including the independent reviews of the arsenic EA by the SAB and a NDWAC working group).

Finally, along with disaggregating benefit and cost (and incremental net benefit) information by system size (which EPA fails to do), the Agency should also disaggregate the findings according to at least two other dimensions that are highly relevant to the proposed LT2 rule:

- EPA should reveal the costs, benefits, and incremental net benefits of the rulemaking options for filtered systems alone and for unfiltered systems alone. In the current EA, some key portions of the risk analysis are clearly specific to either filtered or unfiltered systems. EPA should continue with its separate analyses (rather than revealing only final benefit and cost outcomes for all systems lumped together) so that the public, stakeholders, and decision-makers can see how much more bang per buck may be derived (if any) from focusing rulemaking efforts on unfiltered versus filtered systems.

- EPA should also isolate the costs and benefits for the provision of the rule related to covering finished water reservoirs. It is useful to enlighten the public and government decision-makers alike as to the benefits, costs, and incremental net benefits of such distinct elements of the proposed rule.

7.3 Conclusions

We have not conducted a detailed review of EPA's cost estimates for the proposed LT2 rule. Nonetheless, it seems unlikely that the information provided regarding confidence intervals is realistic. Many key uncertainties and variabilities appear to have been either ignored or understated to a considerable degree in order to generate a very narrow 90% confidence interval (only +/- 11% around the mean).

A critical disappointment with the Agency's cost and affordability analyses, and with the benefit-cost comparisons portrayed in the EA, is the lack of meaningful disaggregation according to system size. The lumping by EPA of size categories serves only to mask and obscure important information regarding the equity and efficiency implications of the proposed rule. This is a serious flaw and a considerable disservice to the public, stakeholders, and decision-makers.

Finally, the Agency's approach to comparing benefits to costs could be far more meaningful and informative if it also disaggregated the benefit-cost results in two important additional dimensions: (1) filtered and unfiltered system benefits and costs, and (2) the costs and benefits of

the finished water reservoir cover requirement as embodied in the preferred option. The Agency needs to do much better with regard to system size and other levels of disaggregation, and it would take only a modest effort on the Agency's part to do so (if it musters the will).

References

Aboytes, R. and M.W. LeChevallier. 2003. Detection of Infectious Cryptosporidium in Filtered Drinking Water. AWWA Annual Conference.

Aboytes, R., G. Di Giovanni, F. Abrams, J. Nieroda, N. Shaw, R. Kozik, and M.W. LeChevallier. 2000. Advances in monitoring finished water for *Cryptosporidium parvum.* Proceedings from AWWA Water Quality Technology Conference.

Apelberg, B., K. Davidson, L. Deck, A. Hallberg, D. McCubbin, and E. Post. 2003. *Proposed Nonroad Landbased Diesel Engine Rule: Air Quality Estimation, Selected Health and Welfare Benefits Methods, and Benefit Analysis Results.* Prepared for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. April.

Centers for Disease Control and Prevention. 2001. *HIV/AIDS Surveillance Report.* 13(2). Internet: http://www.cdc.gov/hiv/stats/hasr1302c.pdf. Downloaded September 12, 2003.

Centers for Disease Control and Prevention. 2003. Safe Food and Water: A Guide for People with HIV Infection. Internet: http://www.cdc.gov/hiv/pubs/brochure/food.htm. Accessed November 18, 2003.

Corso, P.S., M.H. Kramer, K.A. Blair, D.G. Addiss, J.P. Davis, and A.C. Haddix. 2003. Cost of Illness in the 1993 Waterborne *Cryptosporidium* Outbreak, Milwaukee, Wisconsin. *Emerging Infectious Diseases* 9(4):426-431.

DuPont H.L., C. Chappell, C. Sterling, P. Okhuysen, J. Rose, and W. Jakubowski. 1995. The infectivity of *Cryptosporidium parvum* in healthy volunteers. *New England J. Med.* 332:855-859.

Haas, C.N., C.S. Crockett, J.B. Rose, C.P. Gerba, and A.M. Fazil. 1996. Assessing the risk posed by oocysts in drinking water. *Journal of AWWA* 88(9):131-136.

Hoxie, N.J., J.P. Davis, J.M. Vergeront, R.D. Nashold, and K.A. Blair. 1997. Cryptosporidiosis-Associated Mortality Following a Massive Waterborne Outbreak in Milwaukee, Wisconsin. *American Journal of Public Health* 87(12):2032-2035.

LeChevallier, M.W., G.D. Di Giovanni, J.L. Clancy, Z. Bukhari, S. Bukhari, J.S. Rosen, J. Sobrinho, and M.M. Frey. 2003. Comparison of method 1623 and cell culture-PCR for detection of *Cryptosporidium* spp. in source waters. *Appl. Environ. Microbiol.* 69(2):971-979.

Mead, P.S., L. Slutsker, V. Dietz, L.F. McCaig, J.S. Bresee, C. Shapiro, P.M. Griffin, and R.V. Tauxe. 1999. Food-related illness and death in the United States. *Center for Disease Control and Prevention* 5(5):607-625.

Messner, M.J., C.L. Chappell, and P.C. Okhuysen. 2001. Risk Assessment for *Cryptosporidium*: A Hierarchical Bayesian Analysis of Human Dose Response Data. *Wat. Res.* 35(16):3934-3940.

National Research Council. 1999. Risk Assessment of Radon in Drinking Water. Washington, D.C.: National Academies Press. The number was based on the following study: Ershow, A.G. and F.P. Cantor. 1989. Total Water and Tapwater Intake in the U.S.: Population-based estimates. Bethesda, MD: Federation of American Societies for Experimental Biology, Life Sciences Research Office.

Project Inform. 2003. Preventing Cryptosporidiosis: a Guide for People with Compromised Immune Systems. Internet: http://www.thebody.com/pinf/cryptosporidiosis.html. Accessed September 5, 2003.

Raucher, R.S., M.L. Hagenstad, J. Cotruvo, R. Narasimhan, K. Martin, H. Arora, R. Regunathan, J.A. Drago, and F. Pontius. Forthcoming. Conventional and Unconventional Approaches to Water Service Provision. Awwarf Project #2761. Awwa Research Foundation, Awwa Research Foundation, and American Water Works Association.

Roseberry, A.M., and D.E. Burmaster. 1992. Lognormal Distributions for Water Intake by Children and Adults. *Risk Analysis*, 12(1):99-104.

U.S. Census Bureau. 2003. Census Fact Finder. Internet: http://factfinder.census.gov/servlet/GCTTable?ds_name = DEC_2000_SF1_U&geo_id = 01000US&_box_head_nbr = GCT-PH1-R&format = US-9S. Downloaded September 18.

U.S. EPA. 1998. Regulatory Impact Analysis for Interim Enhanced Surface Water Treatment Rule. U.S. Environmental Protection Agency, Washington, DC.

U.S. EPA. 1999. Estimated per Capita Water Consumption in the United States. U.S. Environmental Protection Agency, Office of Science and Technology, Office of Water, Washington, DC. February.

U.S. EPA. 2000. Estimated Per Capita Water Ingestion in the United States. U.S. EPA Office of Water. April.

U.S. EPA. 2001. Arsenic Rule Benefit Analysis: An SAB Review. U.S. EPA Science Advisory. August.

U.S. EPA. 2003a. Economic Analysis for the Long Term 2 Enhanced Surface Water Treatment Rule. Prepared for U.S. EPA by Cadmus Group. June.

U.S. EPA. 2003b. Federal Register. 40 CFR Parts 141 and 142, National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rules; Proposed Rule. 68(154).

U.S. EPA. 2003c. Analysis and Findings of the Gallup Organization's Drinking Water Customer Satisfaction Survey. U.S. Environmental Protection Agency, Office of Groundwater and Drinking Water. August 6.

Viscusi, W.K. and J.E. Aldy. 2003. The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World. Journal of Risk and Uncertainty, 27:1; 5-76.

WQA (Water Quality Association). 1999. 1999 National Consumer Water Quality Survey. Lisle, Ill.: WQA.

A. Issues with EPA's Modeling of *Cryptosporidium* Occurrence[1]

A.1 Summary of *Cryptosporidium* Modeling

The analyses performed with the objective of estimating the national distribution of *Cryptosporidium* demonstrate that if you are certain of the answer before you begin the analysis you can certainly fine-tune your model so that your preconceived notions are realized. This effort has some serious flaws that are best described as indefensible assumptions. The underlying belief in all of these assumptions is that *Cryptosporidium* is ubiquitous in the environment and that when samples are taken there are nearly no zero counts that are true. Since the majority of the values measured in the ICR for *Cryptosporidium* were zeros (not nondetects), this belief suggests that the method is terrible, but that this fact should not trouble anyone since we already know that *Cryptosporidium* is ubiquitous. We patently disagree with this conclusion and underlying assumption. Throughout the ICR data collection and analysis it has been reiterated that a zero count of a discrete item (an oocyst) is a zero and should not be dealt with in any other way. The fact that a modeling exercise can demonstrate that a zero count could be achieved 70% of the time when there is a concentration less than or equal to 1 oocyst per 10 L does not change the fact that the values counted were zeros and there is no additional information to suggest that any of these zeros were something else.

EPA's overly confident conclusion (based on no data) leads the modelers to a series of other assumptions that we do not trust. EPA has failed to furnish information to reveal the potential sensitivity of the model to these assumptions. The specific assumptions that are of significant concern are the true zero values that have been assumed to be 0.1%. That means that only one in 1,000 draws of the data can be considered true zero values. Experts contacted by us believe this value may be as much as two to three orders of magnitude too low. In other words, as many as one in four samples for which a zero was counted may be a true zero. The authors have confused the question of a true zero in a sample with a true zero in a body of water. This model should not be focusing on the concentration in the source water as a whole since the source waters were not properly sampled to characterize the concentration of *Cryptosporidium* in the water body. The best that can be hoped for with the sampling performed is that a gross estimate of the distribution of *Cryptosporidium* at the intake can be estimated.

1. This appendix was prepared by, and reflects expert insights provided by, Jeff Rosen, Perot Systems, December 2003.

The second value that is of significant concern is the estimate of false positive counts. The modelers used a value of 1%. This value could be much higher especially for the samples in which there are very high counts.

By using the *Cryptosporidium* total counts, EPA has also assumed that anything that looks like an oocyst should be counted. The only oocysts that should have been included in these analyses are those that could be identified with high certainty. This would limit the counts to those oocysts that included recognizable internal structures. Overall, we suggest that a detailed presentation of the sensitivity of the analysis to these three assumptions be presented and considered in the benefit-cost evaluation.

Limits of the ICR data. The collection of the ICR data was done in spite of the assertion at the beginning of the process that the data would not be adequate to develop a national distribution that could be trusted. The issues raised related to a few key issues that, in general, all participants agreed were significant problems. This included (but were not limited to) the poor and highly variable recovery, high possibility of false positives (due to algal cells), inclusion of perceived but empty oocysts in counts, no idea of viability or infectivity, inadequate sample volumes, inadequate number of samples, etc. In spite of all these frailties with the methods and the expected results, EPA proceeded with the collection of the data. When the data came back and were clearly inadequate for the intended purpose, rather than admitting that the data were poor and could not be used to reliably to define a national distribution of occurrence of *Cryptosporidium* in source waters, EPA chose to use a very complex modeling approach with a number of unvalidated assumptions to define a national distribution in spite of all the frailties. Assumptions used in the analysis presented in Appendix B of the EA, but which are not defended, include:

The assumed false positive rate (page B-9). It is the opinion of experts that the false positive rate is analyst dependent and in fact could be analyst dependent along with further unquantifiable factors like the hour of the day, the amount of sleep, the expertise of the analyst, etc. It is very possible that some of the very high concentrations that were detected were mostly false positives. After all, if an analyst decided that one algal cell was an empty oocyst, then perhaps the same analyst decided that 85 of the algal cells were oocysts. The high end of the distributions is based on a very small percentage of the samples collected. If these numbers are wrong then the higher bins are not properly defined.

A zero count is once again being called a nondetect (for example, see page 3-19). AWWA has raised this issue repeatedly with EPA, and it is very disconcerting to once again see the zero counts being referred to repeatedly as “nondetects.” As stated in the EA at Section 4.1.2.1, “The model is not limited by the observed detection limits and predicts a positive concentration for nearly all the observed non detects.”

This indicates that the modeling effort continues to insist that a zero is not a zero even though experts assert that some significant percentage (as high as 50%) of the zeros probably indicate that there are no oocysts in the sample. This says nothing about whether or not there are any oocysts in the body of water. If the intent of this effort is to characterize the concentration of oocysts in the water bodies, then the monitoring should be designed to properly assess the concentration in the water body. Such a study could be (but was not) designed. Disappointment in the quality of the data collected was predicted and anticipated before the sampling was required by the ICR.

This issue continues to plague the analysis on page B-12 of the EA. Here the assertion that almost none of the observed values are really zeros is established in the nomenclature of the model. The approach taken is counter to the advice presented by experts and counter to the discussions that we had during the technical working groups. Specifically, the definition of the true zero probability, which is called Z_i, should be the probability that an observed zero (not a nondetect) is in fact really a zero. This should be defined as a zero in the sample taken, not the absence of pathogens from the source body of water. The sampling unit here is a sample taken near the intake, not the water body. There is no way to estimate the probability that there is not a single oocyst in a body of water (read Lake Superior) from which the sample was taken. In fact, the *crypto* sample taken does not target quantifying the oocyst concentration in the water body. No scientist in their right mind would suggest that a single 10-100 L sample taken on one day at a single location in a large body of water could be used to characterize the concentration of *Cryptosporidium* in the entire water body. Instead they would suggest multiple samples at multiple locations at multiple time periods.

So if we assume that the sampling was meant to characterize the water near the intake, then the model should not be taking the leap from the one sample to the concentration of the entire water body. Returning to the true zero value then, we should estimate the number of samples taken for which a zero count was observed and ask the experts what is the probability that the sample taken and counted as zero was truly a zero value. The answer obtained from one expert was between 25 and 50%. This is a probability that should be denoted by Z_{ij}, not Z_i, to reflect the probability from sample j taken at location i.

It can be argued that this original intent is demonstrated on EA page B-12, where the authors mention that the true zero range explored during model development, evaluation, and validation was between 0 and 50% (page B-12, 4th paragraph). After exploring ranges from 0 to 50%, EPA chose to set their value for Z_i to 0.001% which is four orders of magnitude lower than one expert thinks it should be. With Z_i set to 0.001%, one out of every 100,000 values of Z_i generated would be a 1, there by making the concentration Z_{ij} equal to 0. There is no justification for this value nor the philosophy under which the model is defensible.

Z_i is presented as a probability, which it is not. In fact, it is either a 0 or a 1. Since the multiplier in the model is $(1 - Z_i)$, we assume that once out of every 100,000 draws of Z_i a 1 will be drawn and the concentration will be defined as a zero. A number of experts were asked if this is reasonable, and they have indicated that it is not. EPA thus needs to explain the following:

1. What exactly is Z_i and how was it decided to apply the true value to the entire water body instead of to the single concentration?

2. Why is a true zero being applied for each water body instead of estimating it for each sample?

3. What is the effect of using a true zero value on each concentration instead of for the water body?

Likewise, EPA needs to show the sensitivity analyses related to values of true zero from 0.001% to 50%. Does it matter? How will it change the results? What effect will it have on the benefit-cost analyses?

This same issue continues to emerge as a pivotal matter in yet another location in the document. In Section B.1.3 (page B-5 — the basis for modeling), the authors again try to defend that a counted zero is something other than a zero. The discussion is missing one critical phrase in the second paragraph. It says ". . . and if the "true" underlying average concentration in the source water is 0.1 oocyst per liter, it is expected from the Poisson distribution that no Cryptosporidium oocysts would be observed in approximately 74 percent of the samples." Since the chart shown in Exhibit B.1 is a cumulative frequency distribution, this statement is missing five critical words "less than or equal to" It should read ". . . and if the "true" underlying average concentration in the source water is ***less than or equal to*** 0.1 oocyst per liter, it is expected from the poisson distribution that no Cryptosporidium oocysts would be observed in approximately 74 percent of the samples." The true value could be 1 oocyst per liter or it could be 1 oocyst in 10 liters or 1 oocyst in 100 liters. The fact of the matter is that the sampling that was done counted 0 not 1. The method and sampling plan is targeted at counting discrete oocysts. While EPA's arguments are elegant, they are irrelevant. At the end of the argument, we still do not know if the value is a 1 or a zero. We counted zero. We discussed the counting of zero with many statistical experts who all affirmed that the answer was that when a zero is counted with discrete items, it is a zero and not some number less than a concentration that you don't feel you can detect because of the limitations in the method. It is irrelevant if a model can demonstrate that with sufficient assumptions it is possible to say that the answer could have been as high as 1 oocyst per liter. It was not. If this was a level that was needed to answer the question asked, then the monitoring should have been designed to have a larger volume or a larger sample size. The expected count can be manipulated based on models and assumptions, but in the end, the count (not the expected count) is the value that matters.

B. Issues and Practices for Valuing Time

EPA's "Enhanced COI" is an attempt to account for the value of time "lost" by individuals when they are sick. More precisely, it is an attempt to reflect the utility (individual welfare) lost when an individual — due to illness — cannot use their time as productively or in the manner they would otherwise prefer. In other words, the intent is to reflect the loss of well-being associated with having some daily activities impaired or restricted because of illness.

There are several problems and inconsistencies relative to standard and best practices in economics with how EPA develops and applies its Enhanced COI approach:

- The approach can be simplified, because there is no need to make a distinction in the discussion between "leisure time" and "unpaid work time" (household production). When away from work, an individual will choose the activity that has the highest value (yields the most utility), regardless of what it is. Overall, EPA is searching for a proxy for the value of time away from work.

- In recreation demand modeling, this same issue of valuing time away from work arises. For example, in using travel cost-based models developed to value recreation, the "cost" of time spent en route to a recreation site (and perhaps the cost of time at the site while recreating) might be considered as part of the incurred cost of the experience, and thus could be used in estimating a demand curve and associated recreational "values." In recreation studies (for which a vast peer reviewed literature exists), time in transit is predominantly valued at some fraction of the wage rate, such as one-third (and on-site time is rarely included in the valuation).

- Why is a fraction of the wage rate appropriate (rather than 100% of the pre- or post-tax wage rate)? From recreation studies, the following logic has been well articulated:

    - Many workers are salaried, and they do not have an opportunity to make more money at their current jobs by working longer hours.

    - If a worker were to take on a second job, it is likely that the second job would pay less, per hour, than the primary job. This is true for part-time work in many cases, which offers limited benefits and often lower hourly wages than a full-time job.

    - The primary job is more likely to have higher wages than a second job simply because a worker is more likely to take a job with higher wages than lower wages, all else constant. The worker will typically take the most lucrative position available given the level of training and education.

- There have been arguments that the value of time should be greater than 100%, if the worker is putting in overtime hours and earning wages higher than the base wage (e.g., time-and-a-half holiday work). However, this type of work is often sporadic and unpredictable for the worker and is limited in availability, and the quantity is not controlled by the worker. Therefore, economists use the previous arguments and use an estimate less than 100% of the wage rate.

- Finally, it is unreasonable to assume that the individual could be doing professional work during all waking hours.

- The recreation demand literature is one good source of support for the use of one-third the wage rate as the opportunity cost of time away from work. Some relevant literature includes:

  - Carson, R. W.M. Hanemann, M. Costanzo, and T. Wegge. 1991. *1991 Southeast Alaska Sport Fishing Economic Study*. Prepared for the Sport Fish Division, Alaska Department of Fish and Game. Sacramento CA: Jones and Stokes Associates, Inc.

  - Morey, E.R., W.S. Breffle, and P.A. Greene. 2001. Two nested constant-elasticity-of-substitution models of recreational participation and site choice: an "alternatives" model and an "expenditures" model. *American Journal of Agricultural Economics* 83(2):414-427.

  - Needelman, M.S. and M.J. Kealy. 1995. Recreational swimming benefits of New Hampshire lake water quality policies: an application of a repeated discrete choice model. *Agricultural and Resource Economics Review* 24(1):78-87

  - Parsons, G.R. and M.J. Kealy. 1992. Randomly drawn opportunity sets in a random utility model of lake recreation. *Land Economics* 68(1):93-106.