5. CASE STUDY OVERVIEW AND PARAMETER ASSUMPTIONS

The case study presented in Chapter 6 demonstrates how a cost-effectiveness analysis could be developed to evaluate alternative drinking water disinfection technologies. This chapter first provides an overview of the case study, including the technologies evaluated and the health effects considered (Section 5.1.). It then provides details on the data, assumptions and calculations used to determine the central tendency valves and distributions for many of the case study parameters (Sections 5.2. through 5.8.). Refer to Figure 4-3 for a graphical representation of the key components in applying the framework in this case study.

5.1. CASE STUDY OVERVIEW

The treatment and disinfection of drinking water is a primary public health intervention and prevention measure. Under the rules and regulations developed by U.S. EPA, state, and local authorities, decisions regarding how drinking water is treated are made by local purveyors for the community or communities they serve. The treatment decisions faced by purveyors include the consideration of the potentially countervailing risks posed by infectious pathogens and D/DBPs. Although other potentially harmful substances may also be in drinking water (e.g., heavy metals and pesticides), their presence does not bear on decisions regarding disinfection options, the subject of this case study. They are therefore not considered in this analysis.

This limited case study is developed for a hypothetical water distribution system. Three treatment options are evaluated in terms of their impact on microbial risks (GI illnesses and mortality), their impact on DBP-induced risks (cancer, reproductive toxicity, and developmental toxicity), and their financial costs. Specifically, the baseline treatment technology (a standard

treatment train of coagulation, sedimentation, sand filtration and chlorine disinfection) is compared to the baseline treatment technology augmented by either of two supplemental technologies. The first supplemental technology, which benefits all tap water consumers, is the addition of ozone pretreatment to the baseline treatment train. This treatment increases the fraction of pathogens inactivated. Ozone pretreatment decreases the concentration of many DBPs but increases others, in particular, brominated compounds. The second supplemental technology is the installation of point-of-use water filters in the homes of individuals with compromised immune systems. The benefits of this technology are, of course, limited to the consumers receiving the filters. It is important to note that in this limited case study, all three technologies operate in a steady-state environment. That is, it is assumed that no malfunctions occur that cause performance to deviate from system specifications.

The remainder of this section describes the alternative treatment technologies in greater detail (Section 5.1.1.) and the health risks considered in the case study (Section 5.1.2.). It then provides a conceptual overview of the case study's quantification of the health consequences for each technology and the financial costs addressed by the case study (Section 5.1.3.).

5.1.1. Alternative Treatment Technologies. Figures 1-1 and 1-2 illustrate the baseline technology and the baseline technology with the addition of ozone, respectively. The baseline technology is typical of current practice for water purveyors in the United States. The first supplemental technology consists of the addition of ozone pretreatment prior to the coagulation phase of the standard treatment train. Ozone pretreatment is a reasonably feasible alternative approach for which information on microbial treatment efficiency and DBP production are available. Ozone pretreatment is considered here because of its potential effectiveness at

inactivating *Cryptosporidium*, the pathogen featured in this case study. Both treatment alternatives are identical in all respects except for the addition of ozone.

The baseline technology is also compared to its use along with the installation of in-home filters in the dwellings of individuals with compromised immune systems. The AIDS subpopulation serves as a proxy for this group. For the purpose of case study, these filters are assumed to completely remove all microbial agents and to have no effect on DBP concentrations.

5.1.2. Health Risks Considered. DBP-induced health effects are assumed to include cancer, developmental toxicity, and reproductive toxicity. As is typical, the case study assumes that individual DBP cancer risks are a linear function of the average daily dose. As described in Section 5.3., a response addition model was used to estimate mixtures risk for this endpoint. For the purpose of assigning an economic cost to each expected case of cancer, a factor that depends in part on its latency, it is assumed that DBP-induced cancer is manifest as bladder cancer, colon cancer, and cancer of the rectum.

The response addition model was assumed to estimate mixtures risk for the reproductive endpoint. In general, reproductive toxicity may be either reversible (infertility) or permanent (sterility) (U.S. EPA, Guidelines for Reproductive Toxicity Risk Assessment, 1996). Because of the compensatory nature of the reproductive system, however, DBP-associated effects are assumed to be reversible. More specifically, it is assumed that infertility depends only on current DBP exposure. Finally, it is assumed that both males and females may suffer DBP-induced infertility.

The response addition model has also been assumed to estimate mixtures risk for the developmental endpoint. Developmental effects have been commonly considered to be

represented by those events resulting from a post-conception, prenatal exposure, and are thought to be manifest at the time of birth. While embryo and fetal deaths are considered by the Agency as one expression of developmental toxicity, the loss of a fetus here was not addressed. Moreover, developmental effects may conceivably include relatively minor defects (e.g., the loss of hearing in one ear). For the purpose of this assessment, developmental effects are assumed to be permanent, and to result in severe lifetime dependency and a decreased life expectancy. This assumption almost certainly overstates the impact costs associated with DBP-induced developmental toxicity since it places an exaggerated value on each event's cost. The need to revise this assumption so that it is more realistic can be assessed as a part of the sensitivity analysis described in the case study.

Pathogens considered in this analysis are limited to *Cryptosporidium parvum* for three reasons. First, the technologies considered in this case study have a differential effect on the concentration of viable *Cryptosporidium* oocysts in tap water. This difference reflects, in part, the fact that *Cryptosporidium* oocysts are more resistant to chlorine disinfection and are smaller than many other protozoan cysts. Therefore, *Cryptosporidium* oocysts are more likely to pass through traditional water treatment disinfection and filtration processes. Removal of other pathogens by routine water treatment is far more effective, leaving much smaller potential benefits to be accrued by improved disinfection. Second, *Cryptosporidium* is responsible for a great deal of the morbidity and mortality related to drinking water consumption; cryptosporidiosis can occur as a severe and protracted illness and can cause death in immunocompromised individuals such as a person with AIDS (Flanigan et al., 1992, McGowan et al., 1993). There are currently no

effective therapeutic agents to treat Cryptosporidiosis. Finally, the ozone pretreatment system has a non-trivial impact on the concentration of *Cryptosporidium* in finished tap water.

The assumed concentration of *Cryptosporidium* in source waters for the community treatment plant was derived from measurements taken at the intake for the treatment plant in Trenton, N.J. While the efficacy of *Cryptosporidium* removal was assumed to be 100% for the point-of-use devices, the estimated efficacy of the baseline treatment and the ozone-supplemented treatment were both assumed to be less than 100%, based on studies conducted at the pilot facility operated by U.S. EPA. Risks of contracting diarrhea and related sequelae, such as severe illness and death, were estimated from the predicted exposures.

5.1.3. Quantification of Health Consequences and Financial Costs. Health consequences, which are also referred to as the "health costs" associated with a treatment technology, are the health effects resulting from either the presence of infectious agents or DBPs in drinking water. Health effect costs depend on three factors: the tap water consumption rate, the incremental probability of an adverse health effect associated with each liter of water consumed, and the cost (measured in lost QALYs) associated with each health event.

Section 5.2. quantifies the tap water consumption rate. DBP-induced health risks are assumed to depend on the total tap water consumption rate since it is assumed that the concentration of DBPs is unaffected by actions taken after the water leaves the tap (e.g., by heating the water). On the other hand, it is assumed that heating tap water inactivates microbial agents. Section 5.2. therefore quantifies the consumption of unheated tap water, as well.

Sections 5.3. through 5.5. quantify the incremental risks associated with tap water consumption. Section 5.3. quantifies DBP tap water concentrations and DBP slope factors,

Section 5.4. quantifies *Cryptosporidium* concentrations in tap water, and Section 5.5. quantifies the probability of infection, illness, and mortality associated with *Cryptosporidium* exposure.

Section 5.8. quantifies the lost QALYs associated with each of the health endpoints considered in this analysis. These costs depend, in part, on the social discount rate, the value of which is detailed in Section 5.7.

Financial costs considered by the case study are limited to the direct costs of implementing the technologies evaluated. These costs, which are detailed in Section 5.6., consist of the capital costs necessary for installing the technology, and the ongoing operational costs. Other costs, such as medical treatment costs and the lost productivity costs stemming from morbidity and mortality, could also be included in a more expanded analysis. Limiting costs considered to direct technology costs may be realistic if it is assumed the analysis is to be used primarily by tap water purveyors.

5.2. TAP WATER CONSUMPTION

Risks associated with exposure to tap water depend on the quantity of tap water consumed daily. In the case of DBP-induced risks (cancer, reproductive toxicity, and developmental toxicity), it is assumed that risk is proportional to total tap water consumption measured in L/kg-day. Section 5.2.1. quantifies this rate for the general population. In the case of microbial risks, it is assumed that risk depends on the consumption of unheated tap water measured in L/day. Section 5.2.2. addresses this consumption rate. Section 5.2.3. describes necessary adjustments to these results to reflect tap water consumption among the AIDS subpopulation.

5.2.1. Total Tap Water Consumption (L/kg-day). Table 3-6 in U.S. EPA's Exposure Factors Handbook (U.S. EPA, 1997) quantifies total tap water consumption in mL/kg-day for individuals of all ages. Table 5-1 reproduces these rates for the 5th through 95th percentiles of the population.

Table 5-1
Tap Water Consumption in the General Population in mL/kg-day by Age

	Population Percentile						
Age (Years)	**5**	**10**	**25**	**50**	**75**	**90**	**95**
< 0.5	0	0	14.8	37.8	66.1	128.3	155.6
0.5 to 0.9	0	0	15.3	32.2	48.1	69.4	102.9
1 to 3	11.8	17.8	27.2	41.4	60.4	82.1	101.6
4 to 6	10.3	14.9	21.9	33.3	48.7	69.3	81.1
7 to 10	7.4	10.3	16	24	35.5	47.3	55.2
11 to 14	4.9	7.5	11.9	18.1	26.2	35.7	41.9
15 to 19	3.9	5.7	9.6	14.8	21.5	29	35
20 to 44	4.9	7.1	11.2	16.8	23.7	32.2	38.4
45 to 64	8	10.3	14.7	20.2	27.2	35.5	42.1
65 to 74	8.7	10.9	15.1	20.2	27.2	35.2	40.6
> 75	8.8	10.7	15	20.5	27.1	33.9	38.6

Source: Ershow and Cantor (1991) cited in U.S. EPA. 1997

In order to standardize these data for compatible use in the case study, we have taken age-weighted averages[1] of these values to approximate consumption by 5-year increments. The results appear in Table 5-2.

[1]Age-weighted averages represent the average consumption rate for a particular percentile of the population among all individuals in an age group. These values are composite calculations based on data for the appropriate age groups listed in Table 5-1. For example, the age 0 to 4 group values were computed as [0.5 × (age < 0.5) + 0.5 × (age 0.5 to 0.9) + 3 × (age 1 to 3) + 1 × (age 4 to 6)] ÷ 5.

Table 5-2
Tap Water Consumption in the General Population in mL/kg-day by 5-year Age Groups.

Age (Years)	Population Percentile							Arithmetic Mean
	5	10	25	50	75	90	95	
0 to 4	9.14	13.66	23.71	38.5	57.4	82.89	103.03	44.4
5 to 9	8.56	12.14	18.36	27.72	40.78	56.1	65.56	31.2
10 to 14	5.4	8.06	12.72	19.28	28.06	38.02	44.56	21.3
15 to 19	3.9	5.7	9.6	14.8	21.5	29	35	16.3
20 to 24	4.9	7.1	11.2	16.8	23.7	32.2	38.4	18.3
25 to 29	4.9	7.1	11.2	16.8	23.7	32.2	38.4	18.3
30 to 34	4.9	7.1	11.2	16.8	23.7	32.2	38.4	18.3
35 to 39	4.9	7.1	11.2	16.8	23.7	32.2	38.4	18.3
40 to 44	4.9	7.1	11.2	16.8	23.7	32.2	38.4	18.3
45 to 49	8	10.3	14.7	20.2	27.2	35.5	42.1	21.8
50 to 54	8	10.3	14.7	20.2	27.2	35.5	42.1	21.8
55 to 59	8	10.3	14.7	20.2	27.2	35.5	42.1	21.8
60 to 64	8	10.3	14.7	20.2	27.2	35.5	42.1	21.8
65 to 69	8.7	10.9	15.1	20.2	27.2	35.2	40.6	21.9
70 to 74	8.7	10.9	15.1	20.2	27.2	35.2	40.6	21.9
75 to 79	8.8	10.7	15	20.5	27.1	33.9	38.6	21.4
80 to 84	8.8	10.7	15	20.5	27.1	33.9	38.6	21.4
85 +	8.8	10.7	15	20.5	27.1	33.9	38.6	21.4

Source: Ershow and Cantor (1991) cited in U.S. EPA. 1997

Note: The arithmetic mean value for each age group was computed by fitting a lognormal to the percentile values listed and computing the arithmetic mean corresponding to that distribution's geometric mean and geometric standard deviation.

While these values reflect some sampling uncertainty, it is likely that, given the nature of the survey on which it is based, this uncertainty is not substantial.

5.2.2. Consumption of Unheated Tap Water (L/day). It is assumed that microbial risks stemming from ingestion of tap water reflect only consumption of unheated tap water. That is, it is assumed that typical heating of tap water (e.g., during cooking, or in the preparation of hot beverages) inactivates *Cryptosporidium* oocysts. Table 3-7 in U.S. EPA's Exposure Factors Handbook (U.S. EPA, 1997a) quantifies total tap water consumption in mL/day for individuals of all ages. Table 5-3 reproduces these rates for the 5th through 95th percentiles of the population. The case study does not specifically address the issue of bottled water consumption.

Table 5-3
Tap Water Consumption in the General Population in mL/day by Age Group

Age (Years)	Population Percentile						
	5	10	25	50	75	90	95
< 0.5	0	0	80	240	332	640	800
0.5 to 0.9	0	0	117	268	480	688	764
1 to 3	169	240	374	567	820	1162	1419
4 to 6	204	303	459	660	972	1302	1520
7 to 10	241	318	484	731	1016	1338	1556
11 to 14	244	360	561	838	1196	1621	1924
15 to 19	239	348	587	897	1294	1763	2134
20 to 44	337	483	766	1144	1610	2121	2559
45 to 64	591	745	1057	1439	1898	2451	2870
65 to 74	611	766	1044	1394	1873	2333	2693
> 75	568	728	961	1302	1706	2170	2476

Source: U.S. EPA. 1997

The Canada Department of Health and Welfare (1981) reports both total tap water consumption and unheated tap water consumption for individuals of various ages[2]. These figures appear in Table 5-4.

Table 5-4
Fraction of Tap Water Consumed that is Unheated

	Age (Years)					
	< 3	3 to 5	6 to 17	18 to 34	35 to 54	> 55
Unheated (L/day)	0.46	0.74	0.9	0.68	0.59	0.59
Total (L/day)	0.61	0.86	1.14	1.38	1.55	1.57
Fraction unheated	75%	86%	78%	49%	38%	38%

Source: Canadian Minister of Health and Welfare. 1981.

[2]Beverages included as unheated were drinking water, ice/mix, other types of mixes and reconstituted milk. Categories excluded were tea, coffee, soup, homemade beer and wine, popsicles, and baby formula.

Applying the values in the last row of Table 5-4 to those in Table 5-3 (with appropriate age-based weights) yields the estimates in Table 5-5.

The results in Table 5-5 do not reflect population variability due to differences among individuals in the fraction of tap water consumed that is unheated. Nor does the case study characterize uncertainty in these estimates introduced by the estimated fraction of consumed tap water that is unheated.

5.2.3. Adjustments for Special Subgroups.

5.2.3.1. The AIDS Subpopulation — Perz et al. (1998) reports that members of the AIDS subpopulation "*may exhibit significant avoidance of tap water*." They estimate that unheated tap water consumption among individuals in this subgroup is 70% of that among members of the general population. Perz et al. (1998) do not state whether the total tap water consumption rate for the AIDS subpopulation is also less than it is for the general population. Here, it is assumed that while the consumption of unheated tap water for the AIDS subpopulation is lower than it is for the general population, the total tap water consumption is the same. Specifically, it is assumed that age-specific total tap water consumption rates for the AIDS subpopulation are the same as the corresponding rates for the general population. It is assumed that the age-specific unheated tap water consumption rates for the AIDS subpopulation are 70% of the corresponding rates for the general population. No basis has been identified for quantifying the uncertainty associated with the assumptions underlying this estimate.

5.2.3.2. Pregnant Women — Because pregnant women are at risk for DBP-induced reproductive health effects, and the fetus is at risk for DBP-induced developmental effects, it is important to determine if pregnancy status affects tap water consumption. Data quantifying tap water consumption rates for pregnant women are limited. However, Ershow et al. (1991) report

Table 5-5
Unheated Tap Water Consumption in mL/day

Age (Years)	Population Percentile							Arithmetic Mean
	5	10	25	50	75	90	95	
0 to 4	109	157	258	402	590	837	1007	456
5 to 9	178	246	374	554	788	1044	1216	602
10 to 14	192	277	430	644	915	1234	1460	704
15 to 19	160	233	394	602	868	1182	1431	663
20 to 24	166	238	378	564	794	1046	1262	608
25 to 29	166	238	378	564	794	1046	1262	608
30 to 34	166	238	378	564	794	1046	1262	608
35 to 39	128	184	292	436	613	808	975	470
40 to 44	128	184	292	436	613	808	975	470
45 to 49	225	284	403	548	723	934	1093	582
50 to 54	225	284	403	548	723	934	1093	582
55 to 59	222	280	397	541	714	922	1079	574
60 to 64	222	280	397	541	714	922	1079	574
65 to 69	230	288	393	524	704	877	1013	559
70 to 74	230	288	393	524	704	877	1013	559
75 to 79	214	274	361	490	641	816	931	517
80 to 84	214	274	361	490	641	816	931	517
85 +	214	274	361	490	641	816	931	517

Source: Canadian Minister of Health and Welfare. 1981.

Notes: The arithmetic mean value for each age group was computed by fitting a lognormal to the percentile values listed and computing the arithmetic mean corresponding to that distribution's geometric mean and geometric standard deviation.

that the 50th percentile consumption rate among pregnant women is 1.1 L/day, while the 90th percentile daily consumption rate is 2.2 L/day. These values are very close to the 50th and 90th percentile consumption rates, respectively, for members of the general population ages 20 to 44 year old, as reported by U.S. EPA (1997a, Table 3-7). As detailed in Table 5-3 of this report, those rates are 1.144 L/day and 2.121 L/day, respectively. The case study therefore makes no adjustments to the tap water consumption rate to reflect potential differences between pregnant women and the general population.

5.3. DBP RISKS

This section describes the computation of risks associated with exposure to DBPs in drinking water. It is important to note here that although the DBP-induced risks are of concern, results of the case study show that the DBP risks contribute far less to the CEA results than do the microbial risks; thus, the magnitude of the DBP risks and their attendant uncertainty are relatively less important to this particular analysis (see Chapter 7 on interpretation of the case study). Under other treatment scenarios, however, DBP risks could become a more prominent factor in influencing the results.

It is important to keep in mind that the goal here is to make reasonable estimates of human health risks that reflect changes in the DBPs that are produced and in their concentrations and that are comparable across different drinking water treatment types and source water characteristics. An examination of the epidemiologic literature suggests that cancer, reproductive and developmental endpoints are the human health effects of concern in the drinking water; thus these effects need to be reflected in the case study analysis, even if the risk estimates themselves are highly uncertain. Several factors contribute to the uncertainty of estimating risks associated with DBPs: stochastic uncertainty in bioassay data; extrapolation of animal-derived toxicity values to humans; variation in the presence and concentrations of DBPs in the drinking water, seasonal variations in source water conditions, the presence of sometimes large amounts of unidentified halo-organic materials, variations in drinking water intake, and the assumptions that are made as the basis for estimating the mixtures risk.

In the case study, response addition is assumed as a component-based method for joining dose-response and exposure data to estimate cancer, reproductive and developmental risks from exposure to the complex mixture. As stated below (Section 5.3.1.), response addition carries

with it an assumption that the components of the mixture are considered to be functionally independent of one another at low exposure levels; a similar mode of action or similar effects across chemicals are not required (Mumtaz and Hertzberg, 1993). Response addition has often been assumed to estimate cancer risks for a mixture, but is not generally used for noncancer risks because of the assumed existence of component toxicity thresholds. However, for all endpoints, it is possible that a mixtures toxicity threshold exists that would potentially be lower than any of the individual component thresholds, such that estimation of mixtures risk at these individual subthreshold dose levels is reasonable. Response addition works well for this problem because of the need to compare DBP-induced risks for the endpoints of concern at extremely low environmental exposures. This procedure, however, is being used to demonstrate the CRFM and is not a peer-reviewed, accepted method for DBP risk estimation.

Other approaches could certainly be taken, each associated with its own set of assumptions and limitations. Dose-addition is generally preferred for noncancer endpoints; an assumption is required of similar mode of action across all chemical components of the mixture. Dose-addition would be another reasonable choice for the noncancer endpoints as it also addresses the issue of a mixtures toxicity threshold. Proportional- response addition is a hybrid of dose addition and response addition, where risk is estimated for individual components at the total mixture dose and then scaled back by the proportion of the component in the mixture; this approach requires similar effects across chemicals. Another approach is to use human cancer, reproductive or developmental data from the epidemiologic literature. However, the extant epidemiologic data do not distinguish the risks across various treatment technologies and are, therefore, not useful for estimating health risks across specific treatment trains and source waters. A final approach is to develop toxicity data directly on drinking water mixtures or similar

mixtures that represent specific treatment trains and source water. To date, these data are not available for mixtures risk assessment.

Section 5.3.1. describes the response addition model used to calculate the incremental mixtures risk of disease for the known components, summing across chemicals, as the product of a slope factor (representing the potency of each DBP compound), the concentration of each DBP in tap water, and the tap water consumption rate. Section 5.3.2. details the assumptions made for tap water consumption as it relates to DBP exposure. Section 5.3.3. describes the use of data published by Miltner et al. (1990) to quantify the concentrations of identified DBPs and of unidentified Total Organic Halides (TOX). Section 5.3.3. also shows how this analysis estimated the concentration of Bromate in tap water, which Miltner et al. did not measure. Section 5.3.4. summarizes the toxicologic data and modeling efforts used to estimate the slope factors for individual DBPs. Section 5.3.5. describes how these data were used to characterize input distributions for the slope factors, concentrations and tap water consumption values. In the case study, these are repeatedly sampled and the values multiplied in order to estimate health risks for cancer, developmental toxicity, and reproductive toxicity from exposure to both known and unknown DBPs.

5.3.1. Response Addition Model. For this case study, response addition was assumed across chemicals to estimate cancer, developmental, and reproductive human health risks from exposure to a mixture of DBPs found in the distributed drinking water. The response addition model assumes that the components of the mixture are functionally independent of one another at low exposure levels (Mumtaz and Hertzberg, 1993), so that the component exposure risks at low concentrations may be added together (see Appendix A-5 for a more detailed explanation). Because the response addition model is not constrained by the requirement of a similar mode of

action across the chemicals in the mixture, it allows for combining risks across different types of endpoints. Response addition is particularly useful when the effects of concern are thought to be present at low dose levels for each of the component chemicals, even though they are highly unlikely to be observable at these low levels in the environment; the mixture risk is then the sum of the individually low risks of the independently acting component chemicals. For example, response addition has often been used for the risk assessment of mixtures of carcinogens (Gaylor et al., 1997; U.S. EPA, 1989). Equation 5-1 describes the response addition model used in the case study.

$$R_m = \left(1 + \frac{C_u}{C_k}\right) \cdot Y\left(\frac{L}{kg - d}\right) \cdot 0.001\left(\frac{mg}{mg}\right) \cdot \sum_{i=1}^{n} S_i\left(\frac{1}{mg / kg - d}\right) \cdot C_i\left(\frac{mg}{L}\right) \qquad (5\text{-}1)$$

where:

n = the number of DBPs in the mixture known to cause a specific effect R_m = the total mixtures risk from both the n known DBPs and the unidentified Total Organic Halides (TOX) in the mixture

C_i = the concentration (μg/L) of the i^{th} DBP in the mixture

C_u = the total concentration (μg/L) of the measured but unidentified TOX in the mixture associated with causing a specific effect

C_k = the sum of the concentrations (μg/L) of the n DBPs in the mixture known to cause a specific effect

Y = the daily human tap water consumption per body weight (L/kg-d)

S_i = the slope factor for humans $(mg/kg\text{-}d)^{-1}$ for the i^{th} DBP

The final risk estimate, R_m, for the entire mixture is a combination of risks from the DBPs known to cause a specific effect and from the unidentified TOX associated with that effect. This total risk is reflected by the first term on the left side of equation 5-1, $(1 + C_u/C_k)$. The term, $(1 + C_u/C_k)$, is multiplied by the total risk for the known DBPs, which is estimated as the product of the other terms on the right side of equation 5-1. This calculation effectively sums the risk from

the known DBPs and a scaled value of that known risk equal to the ratio of the concentration of unidentified TOX to the total concentration of known DBPs (C_u/C_k). To characterize the distribution of plausible values for R_m, distributions were developed for each of the parameters, C_i, C_u, S_i, and Y (see Section 5.3.5.).

5.3.2. Tap Water Consumption (Y). For DBP risk estimation, the daily human tap water consumption adjusted for body weight (L/kg-d) is used in equation 5-1 to estimate the DBP exposures (see Section 5.2.). Total tap water consumption is used (rather than *unheated* tap water consumption, as in the case of microbial risks) because it is assumed that heating does not substantially affect DBP concentrations. This assumption may lead to the slight overestimation of risk since it is known that the more volatile DBPs, specifically, the trihalomethanes, will be removed by heating. On the other hand, many other DBPs, such as the acids, will remain in the water. Because of a lack of data specifically addressing the removal of these DBPs from water by heating, no adjustment is made for this phenomenon. It is also recognized that the inhalation pathway is a potential route of exposure for the more volatile DBPs (e.g., Jo et al., 1990a,b) provide estimates of exposures to chloroform during showering), although this pathway is also omitted from the case study analysis due to a lack of data. Both of these issues are therefore recognized as potential research needs to the extent that DBP risks substantially affect the results of a comparative analysis such as the case study. Results from the case study indicate, however, that it is unlikely that DBP-induced health risks are important compared with microbial risks given the conditions and assumptions of the case study. It is therefore likely that in many comparative analyses of alternative drinking water disinfection technologies, the assumptions used here will be adequate.

5.3.3. Concentration Data (C_i). Concentration data (in μg/L) for individual DBPs in the case study (one with chlorination only, and one with chlorination following pre-ozonation) were adapted from a paper by Miltner et al., (1990) (see Appendix A-5), resulting from a study in which Ohio River water was treated in a pilot plant and then subjected to a simulated distribution system for each of the treatment trains. Table 5-6 lists the resulting concentration data (C_i) used in the case study. These data are slightly different from the Miltner et al. (1990) paper because the means and confidence limits were recalculated from the sampling data assuming a normal distribution and substituting half the detection limit for non-detects instead of zero, which was used in the original publication. The notable exception is that the concentrations for bromate were not sampled at the time of the study and have been estimated (see Section 5.3.3.1.) using more recent information. Estimates were also made for the unidentified TOX in this study (see Section 5.3.3.2.) for use in health risk estimation.

5.3.3.1. Bromate Estimation — Bromate (BrO_3^-) concentrations were not measured in the Miltner et al. (1990) study (see Section 5.3.3. above). In this study, two parallel treatment trains were examined - one with chlorination only, and one with chlorination following pre-ozonation. Under water treatment plant conditions, chlorine will not react with bromide to form bromate. Rather, chlorine reacts with bromide to form bromine, which reacts with organic compounds to form brominated DBPs. Thus, in this study, bromate formation would be realized only when ozone was employed. Ozone reacts with bromide to form hypobromite ion (OBr^-), and ozone reacts with this ion to form bromate (Shukairy et al., 1994).

To estimate the formation of bromate by ozone in the Miltner et al. study, results were employed from two other studies wherein raw Ohio River water was ozonated in the same pilot-

Table 5-6
DBP Concentrations Used in the Case Study
(Adapted from Miltner et al., 1990)

Chemical	Oz Pre-Tmt/Filtration/Post Cl_3 Mean Conc ug/L	Low 95% Conc ug/L	Upp 95% Conc ug/L	No Pre-Tmt/Filtration/Post Cl_3 Mean Conc ug/L	Low 95% Conc ug/L	Upp 95% Conc ug/L
CHCl3	39.55	34.70	44.40	55.50	52.20	58.80
CHBrCl2	21.10	20.90	21.40	24.40	21.90	26.90
CHBr2Cl	13.00	12.20	13.80	10.20	8.80	11.60
CHBr3	1.50	1.10	1.80	0.35	0.00	0.84
CH	5.80	4.90	6.80	4.20	3.60	4.70
MCA	1.46	1.37	1.54	1.44	1.30	1.60
DCA	19.30	18.00	20.60	30.85	28.40	33.30
TCA	10.00	8.90	11.20	20.10	18.60	21.70
MBA	0.28	0.22	0.34	0.29	0.24	0.33
DBA	1.98	1.74	2.20	1.50	1.30	1.70
BCA	6.70	6.50	6.90	8.50	8.30	8.60
DCAN	2.60	2.20	3.00	3.50	2.70	4.20
TCAN	0.05	0.05	0.05	0.20	0.05	0.30
BCAN	1.65	1.44	1.85	1.90	1.50	2.30
DBAN	0.55	0.31	0.78	0.15	0.03	0.27
Bromate	4.00	3.40	4.60	0.00	0.00	0.00
	Unidentified TOX Estimated for Ozone Pre-Treatment*			**Unidentified TOX Estimated for No Ozone Pre-Treatment****		
Devel.	39.00	28.00	50.00	47.00	35.00	58.00
Repro	39.00	28.00	50.00	47.00	35.00	58.00
Cancer	68.00	48.00	87.00	83.00	62.00	104.00

**Total TOX for Ozone = 207; Of this, 59.3% was unaccounted for.*
***Total TOX for No Ozone = 259; Of this 57.5% was unaccounted for.*

Of the unidentified TOX, 32% was estimated to be associated with developmental effects.
Of the unidentified TOX, 32% was estimated to be associated with reproductive effects.

Of the unidentified TOX, 56% was estimated to be associated with carcinogenic effects.
Note: Because of the uncertainties in making these estimates, it is recognized that the estimated risks from exposure to the unidentified TOX could be very broad in range and could conceivably include zero.

scale contactor. Transfer efficiencies, gas/liquid ratios, liquid depths, ozone-to-TOC or DOC ratios, pHs and temperatures were similar to the Miltner et al. (1990) study. In Miltner et al. (1992), the ambient bromide concentration was 37 µg/L. At ozone/TOC ratios below 1 mg/mg, there was no measurable bromate (when the bromate detection level was 7 µg/L). In Shukairy et al. (1994), the ambient bromide concentration was 50.7 µg/L. At an ozone/TOC ratio near 0.8 mg/mg and a dissolved ozone residual near 0.6 mg/L, the bromate concentration was near 4 µg/L.

Thus, the estimate for bromate formation in this study would be near 4 µg/L, which is below the proposed MCL of 10 µg/L. Replication data described in EPA Method 300.1 for bromate suggests that the expected deviation at 4 µg/L would be ± 0.6 µg/L. Table 5-7 describes the basis for the estimate.

5.3.3.2. Unidentified TOX Estimation — In addition to quantifying the concentrations of known DBPs in their study, Miltner et al. (1990) also give a percentage of unidentified (described as "unaccounted for" in the paper) TOX. Through a personal communication with Richard Miltner, the actual quantity of unidentified TOX was estimated; these values appear in Table 5-8.

These unidentified TOX values were then used to estimate the amount that could be associated with producing developmental, reproductive, or carcinogenic health risk (see Tables 5-6 and 5-9). This was achieved using Quantitative Structure Activity Relationship (QSAR) predictions from the computer software program, TOPKAT® (Toxicity Prediction by Komputer-assisted Technology), first introduced in 1987 by Health Designs, Inc. This software program uses statistically based models that are developed from data bases of known chemical toxicity

Table 5-7
Estimated Bromate Formation in Ohio River Water by Ozonation [a]

	Miltner et al., 1990	Miltner et al., 1992	Shukairy et al., 1994
ozone/TOC, mg/mg	0.8	<1	0.81
pH	7.4 - 8.1	7.8 - 8.1	7.4 - 7.65
temperature, ^{o}C	26 - 28	23 - 24	23 - 24
residual ozone, mg/L	0.47	< 0.47	0.6
bromide, ug/L	37 - 50.7 [b]	37	50.7
bromate, ug/L	4 ± 0.6 [c, d]	< 7	4

[a] *all studies utilize same contactor, similar conditions*
[b] *assumed*
[c] *estimated*
[d] *deviation based on replication data presented in EPA method 300.1*

Table 5-8
TOX, ug Cl/L, in Simulated Distribution (stored) Pilot Plant Waters

O_3 / Cl_2		post Cl_2	
207.4 ± 35.4 [a] (± 17%)		258.8 ± 39.2 [a] (± 15%)	
identified = 84.4 (40.7%) [b]	unidentified = 123 (59.3 %) [b]	identified = 110 (42.5 %) [b]	unidentified = 148.8 (57.5 %) [b]

[a] *mean ± std deviation*
[b] *percentages given in Table 6., Miltner et al., AWWA conf proceedings (June 1990).*

Table 5-9

Number of TOPKAT® QSAR Predictions by Endpoint for Known DBPs Not in the Miltner 1990 Sample

Chemical Class	Total Developmental	Total Cancer	Cancer Female Mouse	Cancer Male Mouse	Cancer Female Rat	Cancer Male Rat
Aldyhydes	4	25	22	2	0	5
Acids	33	38	7	4	33	5
Ketones	15	17	4	10	0	9
Lactones	0	8	3	6	1	0
Alcohols	1	4	2	1	1	2
Ethers	3	6	4	4	3	3
Nitriles	9	7	3	0	3	2
Amines	1	2	1	2	0	0
Amides	1	1	0	1	0	1
Halo/Nitro Alkanes and Alkenes	9	25	11	9	8	15
Total	76	133	-	-	-	-
Percent TOX Attributable to Endpoint	32	56	-	-	-	-

data to provide an initial assessment of the toxicity of chemicals lacking in toxicity data, solely from their molecular structures (see Appendix A-5 for additional details). The unidentified TOX reported by Miltner et al. was assumed to consist of 235 known DBPs (i.e., the 253 DBPs identified by Richardson, 1998, minus the 18 DBPs reported by Miltner et al., 1990). Table 5-9 reports the QSAR-based carcinogen and developmental toxicity predictions for these 235 DBPs. The fraction of unidentified TOX associated with carcinogenicity was assumed to equal the fraction of the 235 substances predicted by the QSAR analysis to be carcinogenic; likewise, the

fraction of TOX associated with developmental toxicity was assumed to equal the fraction of the 235 substances predicted by QSAR to be developmental toxicants. Because the TOPKAT® software does not currently have a model to predict potential reproductive toxicity, the developmental toxicity fraction was used as a surrogate to estimate the fraction of unidentified TOX that may be associated with reproductive toxicity.

It is recognized that there is considerable uncertainty in these estimates due to many factors: estimates of TOX made by conversion to the units µg Cl/L; unknowns relative to the actual number of and molecular weights of chemicals that make up the unidentified TOX; toxicity estimates made from a general list of possible DBPs, rather than from lists that are specific to each treatment train under consideration; and possible classifications errors by the TOPKAT® program. This is only one method that could be used for estimating toxicity for unidentified TOX. Increased information relative to the uncertainties listed here or use of additional QSAR models could improve the accuracy of the estimates. Because of these uncertainties, it is recognized that the estimated risks from exposure to the unidentified TOX could be very broad in range and could conceivably include zero.

5.3.3.3 In-Home Water Filtration — The case study presents placement of in-home reverse osmosis treatment systems for drinking water (point-of-use/single tap) in the homes of immunocompromised members of a population as an alternative intervention to the construction of an ozone/chlorine treatment system for an entire population. For this application we assumed that reverse osmosis completely removes *Cryptosporidium* oocysts and does not significantly affect disinfectant and DBP levels in the drinking water.

5.3.4. Toxicologic and Carcinogenic Risk Data. Tables 5-10, 5-11 and 5-12 list the DBPs that are considered to be carcinogenic, developmental toxicants, or reproductive toxicants,

respectively, for use in the case study. These DBPs were selected largely because they have all been identified as DBPs that occur regularly and in relatively large amounts; thus they have been measured in concentration studies, such as the Miltner et al. (1990) study. Among these chemicals, most have been subjected to one or more toxicologic reviews by the Agency and have sufficient quantitative toxicity data available to model their dose-response and to make judgments on their potential contribution to human health risk (see Appendix A-5 for summaries of toxicity information). For the eight carcinogens listed in Table 5-10, the toxicity data bases are strong enough for classification under the U.S. EPA's 1986 Cancer Guidelines as possible (C) or probable (B2) human carcinogens. Although it is recognized that C and B2 classifications do not strongly indicate human carcinogenic potential, nonetheless, in light of the cancer risks suggested by the epidemiologic data from drinking water studies, they are used here as the most likely candidates for providing estimates of human cancer risk that are comparable across treatment alternatives. Likewise, the five haloacetic acids and four haloacetonitriles that are used here to estimate developmental and reproductive risks vary in the strength of the available toxicologic data bases for each, with a particular paucity of information for the reproductive endpoint.

5.3.4.1 Carcinogen Data — Table 5-10 lists the DBPs that are considered to be carcinogenic. For the case study, the oral upper bound slope estimates for the cancer endpoint were taken directly from the Agency's Integrated Risk Information System (IRIS) (U.S. EPA, 1998a) for bromate, chloroform (CHCl3), bromodichloromethane (CHBrCl2), chlorodibromomethane (CHBr2Cl), and bromoform (CHBr3). The upper bound slope estimate for chloral hydrate (CH) is a verified IRIS workgroup value that has not been loaded onto IRIS to date. All of these values were computed for excess risk, using the linearized multistage model that assumes a low dose linear response. The mean slope estimates for these chemicals were

computed by re-running the linearlized multistage model on the IRIS/workgroup data sets and taking the Maximum Likelihood Estimate (MLE) value. Because the body weight conversions for the DBPs on IRIS were based on the assumption of 2/3 power, this assumption was maintained for consistency. The 1996 draft Cancer Guidelines have proposed the use of 3/4 power for the conversion. This assumption could be used in future applications of the CRFM..

The slope estimates for chloroform are not used in the case study because of a recent expert panel cancer assessment (U.S. EPA, 1998b) that employed methodology from the 1996 proposed Cancer Risk Assessment Guidelines (U.S. EPA, 1996a). There is increasing evidence that the carcinogenic mechanism of action for chloroform is not relevant at the low concentrations found in drinking water and that, based on a margin of exposure (MOE) assessment, any concentrations less than 300 µg/l of chloroform are not of concern for human health. Thus, since even the total TOX in the case study is less than this value, chloroform was not used in the cancer risk estimation; concentrations of chloroform were not assumed to interact with other componts of the DBP mixture..

For dichloroacetic acid (DCA) and trichloroacetic acid (TCA), quantitative cancer estimates are not available on IRIS, but qualitative assessments there list B2 and C cancer classifications, respectively. The upper bound and mean (MLE) slope factors for DCA and TCA were back-calculated from risk levels given in Bull and Kopfler (1991), pages 22 and 23, respectively. DCA was also reviewed by the same expert panel as chloroform; the panel indicated that a lack of evidence exists that tumors occur at low doses of DCA in animal studies (U.S. EPA, 1998b); thus it is questionable whether the mechanism of action for cancer is active at the low levels to which humans are exposed. However, the Agency position on DCA falls short of employing the same MOE methodology as was done for chloroform and Agency text (U.S. EPA,

1998b) leaves open the question of low dose mechanism, so DCA was kept in the case study analysis of cancer risk.

5.3.4.2 Developmental and Reproductive Data — Tables 5-11 and 5-12 list the DBPs that are considered to be developmental and reproductive toxicants, respectively. Although several of these chemicals have Reference Doses (RfDs) listed on IRIS, RfDs are endpoint-specific sub-threshold levels and are not useful for the dose-response analysis needed for the case study. Table 5-13 shows chemical names and formulas for these DBPs along with the availability of developmental and reproductive dose-response data for six of the haloacetic acids (MCA, DCA, TCA, MBA, DBA and BCA), four of the haloacetonitriles (DCAN, TCAN, BCAN and DBAN) and one of the trihalomethanes (BDCM) (see Appendix A-5 for more details). Seven of these DBPs (MCA, DCA, TCA, MCA, DCAN, TCAN, BCAN) have been subjects of developmental toxicity studies by a single group of investigators, and three (DCA, MBA, DBA) have been the subjects of male reproductive studies by another group of investigators. These studies were all conducted in rats using gavage administration. The results for developmental toxicity were positive. For reproductive toxicity, the dihalogenated haloacetic acids gave positive results, but the monohalogenated acetic acid (MBA) gave negative results. DBAN, was tested in a short-term developmental and reproductive toxicity screening study in rats by the NTP (1992), with negative results. BDCM was tested in a developmental toxicity screening bioassay with positive results. Adequate developmental toxicity data are lacking for DBA and BCA and for DBAN. A surrogate approach seemed appropriate to fill these data gaps, because the available data indicated that developmental toxicity may be common to the haloacetic acid and

Table 5-13
Availability of Developmental and Reproductive Dose-Response Data

Chemical			Developmental Toxicity[a]	Reproductive Toxicity[a]
Haloacetic Acids				
$ClCH_2COOH$	Monochloroacetic Acid	MCA	y, (+)	
$Cl_2CHCOOH$	Dichloroacetic Acid	DCA	y, +	y, +
Cl_3CCOOH	Trichloroacetic Acid	TCA	y, +	
$BrCH_2COOH$	Monobromoacetic Acid	MBA	y, +	y, -
$Br_2CHCOOH$	Dibromoacetic Acid	DBA		y, +
BrClCHCOOH	Bromochloroacetic Acid	BCA		
Haloacetonitriles				
Cl_2CHCN	Dichloroacetonitrile	DCAN	y, +	
Cl_3CCN	Trichloroacetonitrile	TCAN	y, +	
BrClCHCN	Bromochloroacetonitrile	BCAN	y, +	
Br_2CHCN	Dibromoacetonitrile	DBAN	y,(-)[b]	y, (-)[b]
Trihalomethanes				
$CHBrCl_2$	Bromodichloromethane	BDCM	y, +	

[a]Data are from gavage studies in rats unless otherwise noted.
[b]Data are from a screening-level drinking water study in rats.
y = yes, adequate data available
+ = results were positive for adverse effect
- = results were negative for adverse effect
(+) = results were marginally positive
(-) = results were negative, but a toxicity-based MTD could not be achieved due to taste aversion and consequent refusal to drink higher concentrations of the chemical, and this was a short-term screening study.

haloacetonitrile DBPs. As a provisional measure, DCA was selected as a surrogate for the haloacetic acids and TCAN was selected as a surrogate for the haloacetonitriles.

Dose-response modeling was performed on all possible developmental and reproductive endpoints using human equivalent doses (calculated with a scaling factor of body weight to the 2/3 power to be consistent with the assumption employed for the cancer data on U.S. EPA's IRIS database and in Section 5.3.4.1.) in a linearized multi-stage model with a threshold parameter estimated by the modeling procedure. Note that some of the data are quantal, but other data (body weight, crown-rump length) are continuous and were converted to a quantal measure prior to modeling (see Appendix A-5 for details). Modeling results for all data sets are found in Table 5-14. For many of the data sets, the threshold estimates were above concentration levels for the treatment trains and were therefore not included in any of the risk estimates. This criteria excluded MCA and BDCM entirely from the risk calculations. For the other DBPs, the modeling procedure failed to estimate a threshold value for one or more of the data sets such that the threshold was effectively set to zero. For these cases, scientific judgment was used to look across these data sets for the strongest data set and model results, using factors such as evidence of dose-response in the raw data, larger sample sizes, and adequate goodness-of-fit of the model, to choose a dose-response model. For the calculations of extra risk that were made from these data sets, the estimates made directly from the model were identical to those calculated from the slope factors alone because the low dose region of the dose-response curve is relevant in this context. Therefore, MLE and upper bound slope factors were taken from the modeling results for use in risk estimation.

5.3.5. Input Distributions for Simulation Procedures. For each of the treatment trains and DBP-induced health effects considered in the case study, Monte Carlo techniques were applied to

Table 5-14

Threshold Model Results using $BW^{2/3}$ Scaling Factor*

Data Set [note all are for rats except when noted (DCA, Cicmanec et al.)]	Equiv Human ED_{01} mg/kg-d	Equiv Human ED_{10} mg/kg-d	Threshold mg/kg-d
MCA Smith et al , Fetal body weight	18.9	26.7	11.2
MCA Smith et al , Crown-rump length	15.7	20.2	11.2
MCA Smith et al, Visceral Malformations	16.5	21.8	11.2
DCA Smith et al , Fetal body weight - male	4.7	27.3	2.2
DCA Smith et al , Fetal body weight - female	18.6	40.4	16.3
DCA Smith et al , Crown-rump length - male	5.1	36.2	1.0
DCA Smith et al , Crown-rump length - female	5.1	36.2	1.0
DCA Smith et al , Visceral malformations Total	1.2	12.2	0
DCA Smith et al , Visceral malformations Cardiovascular	1.7	17.6	0
TCA Smith et al , Complete litter resorption	110.5	143.2	106.3
TCA Smith et al , % Postimplantation loss/litter	51.1	88.9	46.8
TCA Smith et al , Fetal body weight - male	0.5	5.2	0
TCA Smith et al , Fetal body weight - female	0.6	6.0	0
TCA Smith et al , Fetal crown-rump length - male	16.2	26.8	15.0
TCA Smith et al , Fetal crown-rump length - female	22.9	37.9	21.4
TCA Smith et al , Visceral malformations Total	25.7	32.2	25.0
TCA Smith et al , Visceral malformations Cardiovascular, total	11.9	23.4	10.7
TCA Smith et al , Visceral malformations Levacardia	1.3	13.8	0
TCA Smith et al , Skeletal malformations	129.7	145.3	128.0
MBA Randall et al., Fetal body weight	4.4	13.7	3.4
MBA Randall et al., Fetal crown-rump length	1.2	12.5	0
MBA Randall et al., Visceral malformations (% affected/litter)	10.2	15.5	6.1
DCA Cicmanec et al , Testicular lesions: degeneration, **dog**	Failed to converge		
DCA Linder et al , Number caput sperm	33.3	74.6	28.8
DCA Linder et al , Number cauda sperm	Failed to converge		
DCA Linder et al., % Motile sperm	12.6	16.5	9.7
DCA Linder et al., Progressive motility	10.8	15.4	9.7

Data Set [note all are for rats except when noted (DCA, Cicmanec et al.)]	Equiv Human ED_{01} mg/kg-d	Equiv Human ED_{10} mg/kg-d	Threshold mg/kg-d
DCA Linder et al., Testicular histopathology: Faulty spermiation	Failed to converge		
DBA Linder et al , Number caput sperm	5.6	7.7	5.4
DBA Linder et al , Number cauda sperm	0.4	4.2	0
DBA Linder et al , % Motile sperm	9.4	13.9	5.4
DBA Linder et al , Progressive motility	9.4	13.9	5.4
DBA Linder et al , Retention Stage IX spermatids per tubule	0.1	1.1	0
DCAN Smith et al , Complete litter resorption	2.4	3.2	2.3
DCAN Smith et al , % Postimplantation loss/litter	2.3	3.6	1.9
DCAN Smith et al , Fetal body weight - male	2.1	4.3	0.8
DCAN Smith et al , Fetal body weight - female	2.6	3.6	2.4
DCAN Smith et al , Fetal Crown-rump length - male	2.8	4.1	2.4
DCAN Smith et al , Fetal Crown-rump length - female	2.3	3.4	2.2
DCAN Smith et al , Visceral malformations Total	1.5	2.3	1.5
DCAN Smith et al , Visceral malformations Cardiovascular	0.2	1.8	0
DCAN Smith et al , Visceral malformations Urogenital	0.9	2.3	0.8
DCAN Smith et al , Skeletal malformations	1.1	3.2	0.8
TCAN Smith et al , Complete litter resorption	0.24	0.97	0.16
TCAN Smith et al , % Postimplantation loss/litter	0.5	1.2	0.4
TCAN Smith et al , Fetal body weight - male	0.2	1.7	0
TCAN Smith et al , Fetal body weight - female	0.1	1.1	0
TCAN Smith et al , Visceral malformations Total	0.05	0.5	0
TCAN Smith et al , Visceral malformations Cardiovascular	0.09	0.9	0
TCAN Smith et al , Visceral malformations Urogenital	0.06	0.7	0
BCAN Christ et al, Complete litter resorption	1.1	3.7	0.8
BCAN Christ et al, % Postimplantation loss/litter	0.6	6.5	0
BCAN Christ et al, Fetal body weight - male	0.8	2.0	0.6
BCAN Christ et al, Fetal body weight - female	1.0	2.8	0.8
BCAN Christ et al, Fetal crown-rump length - male	0.5	4.8	0
BCAN Christ et al, Fetal crown-rump length - female	0.2	1.9	0
BCAN Christ et al, Visceral malformations Total	0.06	0.6	0
BCAN Christ et al, Visceral malformations Cardiovascular	0.07	0.7	0

Data Set [note all are for rats except when noted (DCA, Cicmanec et al.)]	Equiv Human ED_{01} mg/kg-d	Equiv Human ED_{10} mg/kg-d	Threshold mg/kg-d
BCAN Christ et al, Visceral malformations Urogenital	0.5	1.9	0.4
BCAN Christ et al, Skeletal malformations	1.0	3.4	0.8
BDCM Narotsky et al, Complete litter resorption	3.8	6.1	3.5

**See Appendix A-5 for details on dose conversions.*

Equation 5-1 to generate a distribution of human health risk values. The distribution of risks for each DBP-induced health effect was calculated by first randomly drawing concentration, tap water consumption, and slope factor values for each identified DBP.

- The concentration (C_i) of each identified DBP and the concentration of the unidentified TOX (C_u) were assumed to be normal with means and standard deviations consistent with the means and 95th percentile values reported in Table 5-6. These concentration values are specific to the hypothetical treatment plant that is used as a basis for the case study and are based on empirical sampling data. They are not representative of average values for these types of treatment facilities and cannot be used to estimate concentrations where other source waters are used.

- The slope factor (S_i) for each identified DBP was assumed to be lognormal with a geometric mean equal to the 50th percentile and 95th percentile values reported in Tables 5-10, 5-11 and 5-12. Although the confidence intervals reported for slope factors fit using maximum likelihood techniques are theoretically normal, the relative magnitude of the 50th and 95th percentile values the slope factors for many compounds, along with the constraint that the slope factor must be non-negative, indicates that the true confidence intervals must be skewed to the right. The lognormal distribution was used to approximate this skew. The sensitivity analysis reported in Chapter 6 for the case study indicates that DBP slope factor uncertainty did not have an important impact on the case study results, hence indicating that this approximation is sufficient for the purpose at hand.

- The tap water concentrations were sampled from the distributions given in Table 5-2 (Section 5.2.2) which account for differences in body weight and age. As described in Section 5.3.2 above, it is assumed that heating the water does not significantly affect the levels of exposure to the DBPs.

These randomly drawn parameter values were then used in equation 5-1 to calculate a risk estimate, and the process was repeated to generate a distribution of risks. Chapter 6 of this document describes the procedures for these simulations in more detail.

5.4. DISTRIBUTION OF *CRYPTOSPORIDIUM* CONCENTRATIONS IN SOURCE WATER

The concentration of *Cryptosporidium* oocysts in tap water is assumed to equal the source water concentration (Section 5.4.1.) divided by the removal efficiency of the treatment technology (Section 5.4.2.).

5.4.1. Source Water Oocyst Concentrations. The following discussion characterizes the long-term average concentration of *Cryptosporidium* oocysts in source water for a single water treatment plant. The assessment is based on data collected as part of an investigation conducted by LeChevallier et al. (1998). The data consist of 60 quarterly samples and 12 monthly samples taken from water at the intake to the Trenton, NJ Water Works filtration plant.

Table 5-15 summarizes the data collected by LeChevallier et al. (1998). LeChevallier et al. counted the number of inactive oocysts (defined to be oocysts that are either empty or have amorphous internal morphology and presumed noninfectious) and the number of active oocysts (defined to be oocysts with from 1 to 4 possible sporozoites, presumably infectious).

Use of these data to calculate an annual mean concentration and its attendant uncertainty is complicated by three factors. First, it is not clear whether the total (inactive plus active) oocyst count or just the active oocyst count should be used to quantify the *Cryptosporidium* concentration. Second, most of the total oocysts measurements yielded values below the detection limit of 20 per 100L. Third, quality control tests of the sampling method revealed a very low recovery rate, ranging from only 7 to 41% (LeChevallier, 1998). In general, oocyst

Table 5-15
***Cryptosporidium* Concentrations in Source Water at the Intake of the Trenton, NJ Water Works Filtration Plant**

Collection Period	N	Total Oocysts			Active Oocysts		
		N Above Detect Limit	AM[a] for Samples Above Detect Limit	Max	N Positive and Above Detect Limit	AM for Positive Samples Above Detect Limit	Max
Jan to Mar	18	9	42	100	4	19.9	20
Apr to Jun	18	11	58	280	4	40	60
Jul to Sep	18	6	48	140	2	30	40
Oct to Dec	18	3	20	20	2	20	20
All Data	**72**	**29**	**48**	**280**	**12**	**28.3**	**60**

Note: AM is Arithmetic Mean

recoveries and determination of oocyst viability or infectivity have not been reliably measured because of shortcomings in the testing methods commonly used. Clancy et al. (1997) reported that *Cryptosporidium* recoveries ranged from 0 to 138%. Some labs could not recover cysts and oocysts (false negatives) in some of the samples even at the high seeding levels. Methodological difficulties and limitations include the following: low organism recovery; procedures can be time-consuming and require specialized, expensive equipment; nonspecificity of the monoclonal antibodies; determination of viability/infectivity; inability to identify the host of origin; large amounts of algae and debris (many algae autofluoresce resulting in false positives); skill, expertise and training of the microscopist (Schaefer, 1997.)

Total *vs.* active counts: Although dose theoretically depends on the "active" oocyst count, it is also important to use a count for the assessment of exposure that is comparable to the count used for the assessment of infectivity. The infectivity estimate used here (see Section 5.5.) has been developed by Perz et al. (1998) from data collected by DuPont et al. (1995). The

DuPont et al. investigation administered laboratory-prepared *Cryptosporidium* oocysts to 29 volunteers. Because the samples were prepared in the lab, DuPont et al. could control the oocyst species administered and the age of the oocysts (which affects viability). The laboratory setting also offered DuPont et al. an opportunity to assess viability via *in vitro* excystation. By virtue of the experimental design, DuPont et al. were able to report relatively accurate, although still imperfect, estimates of dose expressed in terms of infectious oocysts consumed per day.

In contrast, samples taken from the water intake of the Trenton, NJ plant contained oocysts of unknown age and species. Many of these oocysts are likely not to have been infective because they are either inactive or because their species do not infect humans. For these reasons, it is likely that use of the total oocyst count overstates the infective dose corresponding to the measurement used in the DuPont et al. (1995) study. On the other hand, it is possible that the active oocyst count incorrectly omits oocysts that should be classified as infective. It is therefore possible that use of the active oocyst count understates the infective dose. (It is also possible that the active oocyst count incorrectly includes oocysts that are not infective; hence, use of even the active oocyst count may overstate the infective dose.) For the purpose of the case study, it is assumed that use of the LeChevallier et al. (1998) active oocyst count yields the correct measure of dose (i.e., the measure of dose that best corresponds to the DuPont et al. measure of dose) with 50% probability. Reflecting the possibility that use of the active oocyst count understates the infective dose, it is assumed that with 50% probability, the total oocyst count is the correct measure of dose. The case study finds that its analysis is sensitive to the assumed source water *Cryptosporidium* concentration. Chapter 8 of this report points out that developing improved *Cryptosporidium* data is an important research need.

Non-detects: The non-detect values (and zero concentration values, in the case of the active counts) were replaced with surrogate values by fitting a lognormal distribution to the positive observations above the detection limit. For the total counts, 29 of the 72 observations exceeded the detection limit of 20 oocysts per 100 liters. These 29 values were assigned fractile values of 44/73, 45/73, and so on, up to 72/73, and a lognormal distribution was fit to the data by plotting the log of these values against the normal distribution z-scores corresponding to their ranks. Values from this lognormal were assigned to the 43 non-detects by identifying the 43 values corresponding to the fractiles 1/73, 2/73, and so on, up through 43/73. Figure 5-1 illustrates the plot of the log-transformed measurements above the detection limit against their corresponding z-scores. Although the large number of measurements at exactly the detection limit of 20 oocysts per 100 L appears to be an artifact, the data are reasonably consistent with the straight line in Figure 5-1, thus suggesting that the lognormal adequately describes these measurements.

For the active oocyst counts, only 12 of the 72 measurements exceeded zero. A plot of the log of these values against their normal distribution z-scores appears in Figure 5-2. The same approach was used to calculate surrogate values for the other 60 measurements as was used to calculate surrogate values for non-detects in the total count data set.

For the lognormal distribution fit to the total count data (Figure 5-1), the exponentiated intercept (which equals the geometric mean for this lognormal) is 9.2 oocysts / 100 L, while the exponentiated slope of the line (which equals this lognormal's geometric standard deviation) is 4.0. The arithmetic mean of this distribution is 23.6 oocysts / 100 L. This value does not substantially differ from the arithmetic mean of 25.0 oocysts / 100 L calculated when the non-

Figure 5-1
Log *Cryptosporidium* Concentrations (Total Count) vs. Normal Distribution z-scores for Measurements Exceeding the Detection Limit of 20 Cysts / 100 L

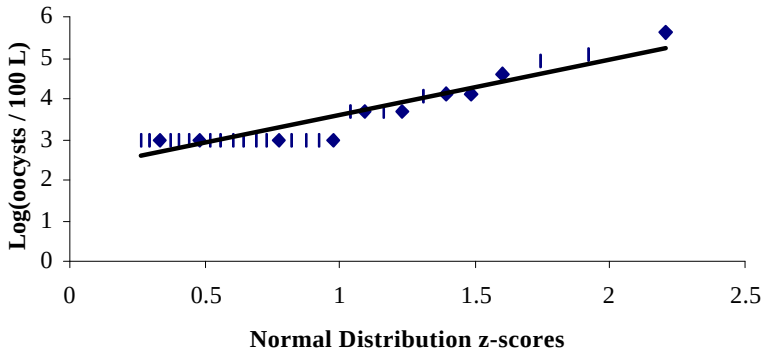
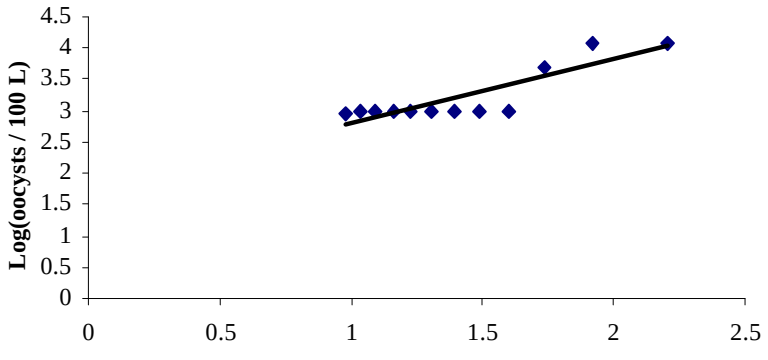


Figure 5-2
Log *Cryptosporidium* Concentrations (Active Count) vs. Normal Distribution z-scores for Positive Measurements



detects are assumed to equal 10 oocysts per 100 L (1/2 the detection limit of 20 oocysts per 100 L).

For the lognormal distribution fit to the active count data (Figure 5-2), the exponentiated intercept (the geometric mean) is 5.8 oocysts / 100 L, while the exponentiated slope of the line (the geometric standard deviation) is 2.8. The arithmetic mean of this distribution is 9.9 oocysts / 100 L. Since there is no defined detection limit for the active count, no comparison can be made to the arithmetic mean calculated by substituting 1/2 of this limit in place of values that are either zero or non-detects.

Table 5-16 details the original values and the calculated proxy values for both the total and active oocyst counts.

Recovery Rates: The authors prepared 20 negative and 20 positive parasite quality control samples for analysis. The recovery rate for the 20 positive control parasite samples ranged from 7 to 41%, with a median of 15% and a geometric mean of 16%. LeChevallier et al. (1998) note that the recovery rate did not appear to exhibit a seasonal trend. Because oocyst concentrations do exhibit a seasonal trend, this finding suggests that the recovery rate does not depend on the true oocyst concentration[3]. Table 5-17 details the recovery rates for the 20 test samples.

It is assumed that each observation in Table 5-16 reflects the impact of any of the recovery rates listed in Table 5-17 with equal probability. Thus, for example, since 1 of the 20 quality control samples yielded a recovery rate of 41%, there is a 5% chance that the 66th total count reported in Table 5-16 of 58.2 oocysts / 100 L represents a true concentration of 58.2 ÷

[3] If the recovery rate did depend on concentration, it would also depend on season because oocyst concentrations vary by season. However, it may also be true that there was insufficient statistical power to detect such an association.

Table 5-16

Trenton, NJ Water Works Filtration Plant Data and Proxy Values Used for Measurements Below the Detection Limit

Percentile	Total Counts		Active Counts	
	Reported Value	Value or Interpolated Value	Reported Value[a]	Value or Interpolated Value
1.4%	<20	0.4	0	0.6
2.7%	<20	0.7	0	0.8
4.1%	<20	0.8	0	1.0
5.5%	<20	1.0	0	1.1
6.8%	<20	1.2	0	1.3
8.2%	<20	1.4	0	1.4
9.6%	<20	1.5	0	1.5
11.0%	<20	1.7	0	1.6
12.3%	<20	1.9	0	1.8
13.7%	<20	2.0	0	1.9
15.1%	<20	2.2	0	2.0
16.4%	<20	2.4	0	2.1
17.8%	<20	2.6	0	2.3
19.2%	<20	2.8	0	2.4
20.5%	<20	3.0	0	2.5
21.9%	<20	3.2	0	2.6
23.3%	<20	3.4	0	2.8
24.7%	<20	3.6	0	2.9
26.0%	<20	3.8	0	3.0
27.4%	<20	4.0	0	3.1
28.8%	<20	4.2	0	3.3
30.1%	<20	4.5	0	3.4
31.5%	<20	4.7	0	3.6
32.9%	<20	5.0	0	3.7
34.2%	<20	5.3	0	3.8
35.6%	<20	5.5	0	4.0
37.0%	<20	5.8	0	4.1
38.4%	<20	6.1	0	4.3
39.7%	<20	6.4	0	4.5
41.1%	<20	6.7	0	4.6
42.5%	<20	7.1	0	4.8
43.8%	<20	7.4	0	5.0
45.2%	<20	7.8	0	5.2
46.6%	<20	8.2	0	5.3
47.9%	<20	8.5	0	5.5

Percentile	Total Counts		Active Counts	
	Reported Value	Value or Interpolated Value	Reported Value	Value or Interpolated Value
49.3%	<20	9.0	0	5.7
50.7%	<20	9.4	0	5.9
52.1%	<20	9.8	0	6.1
53.4%	<20	10.3	0	6.4
54.8%	<20	10.8	0	6.6
56.2%	<20	11.4	0	6.8
57.5%	<20	11.9	0	7.1
58.9%	<20	12.5	0	7.4
60.3%	20.0	20.0	0	7.6
61.6%	20.0	20.0	0	7.9
63.0%	20.0	20.0	0	8.2
64.4%	20.0	20.0	0	8.5
65.8%	20.0	20.0	0	8.9
67.1%	20.0	20.0	0	9.2
68.5%	20.0	20.0	0	9.6
69.9%	20.0	20.0	0	10.0
71.2%	20.0	20.0	0	10.4
72.6%	20.0	20.0	0	10.8
74.0%	20.0	20.0	0	11.3
75.3%	20.0	20.0	0	11.8
76.7%	20.0	20.0	0	12.4
78.1%	20.0	20.0	0	13.0
79.5%	20.0	20.0	0	13.6
80.8%	20.0	20.0	0	14.3
82.2%	20.0	20.0	0	15.1
83.6%	20.0	20.0	19.4	19.4
84.9%	40.0	40.0	20.0	20.0
86.3%	40.0	40.0	20.0	20.0
87.7%	40.0	40.0	20.0	20.0
89.0%	40.0	40.0	20.0	20.0
90.4%	58.2	58.2	20.0	20.0
91.8%	60.0	60.0	20.0	20.0
93.2%	60.0	60.0	20.0	20.0
94.5%	100.0	100.0	20.0	20.0
95.9%	140.0	140.0	40.0	40.0
97.3%	160.0	160.0	60.0	60.0
98.6%	280.0	280.0	60.0	60.0

Notes *A value of 0 in this column means that either the total count was below the detection limit or the total count exceeded the detection limit but the number of active oocysts in the sample was zero.*

Table 5-17
***Cryptosporidium* Recovery Rates for 20 Quality Control Samples**

Percent Recovery	Number of Observations with this Recovery Rate
7	1
9	1
10	1
11	2
13	1
14	2
15	4
17	3
22	1
27	1
31	1
32	1
41	1

41%, or 142 oocysts per 100 L. Likewise, there is a 15% chance that the recovery rate for this sample was 17%, and hence a 15% chance that this measurement represents a true concentration of 58.2 ÷ 17%, or 342 oocysts per 100 L[4].

It is further assumed that the sample size of 72 measurements (72 total counts and 72 active counts) is adequately large to invoke the central limit theorem. Specifically, it is assumed that the average total or active oocyst concentration is approximately normal with a mean equal to the sample mean concentration (corrected for the recovery rate) and a variance equal to the square of the sample mean's standard error. Since the correct recovery rate adjustment for each measurement is not known, there are many possible distributions for the average total or active

[4] It must be noted that it has been assumed that the reported concentration and the recovery rate are statistically independent (as suggested by the lack of a seasonal recovery rate trend, discussed above). However, it is plausible that high reported concentrations are indicative of a high recovery rate. If this is the case, the approach below overstates the uncertainty in the annual average oocyst concentration.

oocyst concentration. Specifically for every possible assignment of the 20 recovery rates in Table 5-17 to the 72 total or active in Table 5-16, it is assumed that there is an equally plausible distribution for the annual average oocyst concentration. The i^{th} such distribution has cumulative distribution function CDF_i, mean m_i and standard deviation s_i. The probability that the average annual total or active oocyst concentration is less than C, conditioned on the assumption that the true sample mean and standard error are, respectively, equal to m_i and s_i, is $CDF_i(C)$ – *i.e.*, area to the left of C under the normal distribution with mean m_i and standard deviation s_i. Hence, the unconditional probability that the average annual oocyst concentration is less than C is the average of $CDF_i(C)$ over all possible distributions, i. Hence, the average of the CDFs defines the distribution of plausible values for the average annual total or active oocyst concentration, reflecting both sampling variability among the measurements in Table 5-16, and uncertainty in the recovery rate applicable to each of these measurements.

The average CDF for the average total or active annual oocyst concentration was estimated by averaging 500 randomly simulated CDFs. This number of iterations yielded an adequately precise characterization of the average CDF. For the total oocyst count, the maximum standard error at any point in the distribution for the sample mean was 0.0074 (the sample mean CDF value at this point was 0.574). For the active oocyst count, the maximum standard error at any point in the distribution for the sample mean was 0.0077 (the sample mean CDF value at this point was 0.536).

Figure 5-3 and 5-4 plot the simulated average CDF values on the vertical axis against normal distribution z-scores on the horizontal axis for the total and active oocyst counts,

respectively. Both plots form nearly a straight lines, indicating that they are well-described by a normal distribution. For the total oocyst count, the best-fit line has an intercept of approximately

Figure 5-3
Average Simulated CDF for the Average Annual Source Water
***Cryptosporidium* Concentration**
(Total Organisms / 100 L)

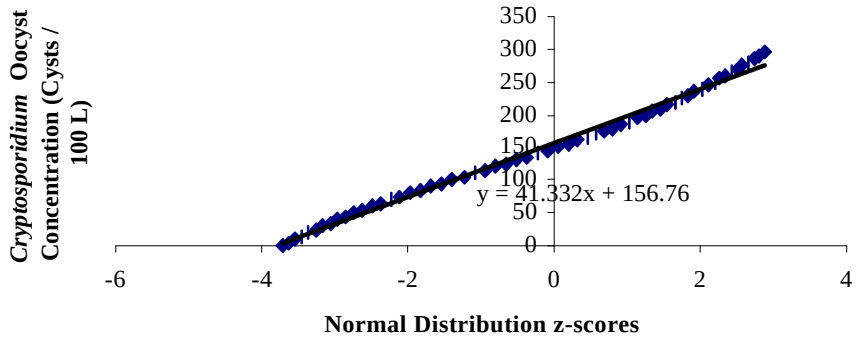
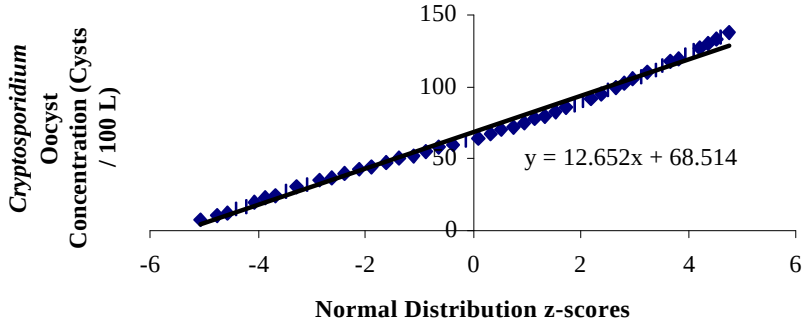


Figure 5-4
Average Simulated CDF for the Average Annual Source Water
***Cryptosporidium* Concentration**
(Active Organisms / 100 L)



156.8 and a slope of 41.3. For the active oocyst count, the best fit line has an intercept of approximately 68.5 and a slope of 12.7. It is therefore assumed that the average annual oocyst concentration is:

- With 50% probability, normal with a mean of 156.8 oocysts / 100 L and a standard deviation of 41.3 oocysts / 100 L; and
- With 50% probability, normal with a mean of 68.5 oocysts / 100 L and a standard deviation of 12.7 oocysts / 100 L.

5.4.2. Fraction of Oocysts Removed by Water Treatment. The concentration of *Cryptosporidium* oocysts in tap water depends on the removal efficiency of the treatment technology (see Table A-1-2, Appendix A1.)

Baseline technology: The removal efficiency of a conventional treatment plant using post-chlorine disinfection is 2 logs (i.e., the concentration of oocysts in the water is reduced by a factor of 10^2, or 100).

Ozone Pretreatment: Ozone pretreatment in addition to the baseline technology reduces oocyst concentrations by between 0.5 and 1.5 logs (a factor of 3.16 to 31.6). The precise removal efficiency is uncertain. Its log is assumed to follow a triangular distribution with a lower bound of 0.5, an upper bound of 1.5, and a mode of 1.0.

In-Home Filters: Intact in-home reverse osmosis filter systems are assumed to remove all protozoa such as *Cryptosporidium* resulting in a finished water free of protozoa. Similar to the operations considered in the two treatment plants, no system malfunctions, etc. are considered in this analysis.

5.5. HEALTH RISKS ASSOCIATED WITH EXPOSURE TO *CRYPTOSPORIDIUM*

Health risks posed by exposure to *Cryptosporidium* in drinking water are modeled by assuming four disease states. These states are based on the model developed by Perz et al. (1998), and the definitions in this case study are based on the definitions proposed in that paper:

- Infection: The detection of oocysts in a subject's feces without the manifestation of any clinical symptoms of illness;
- Mild illness: A single day of diarrhea;
- Moderate to severe illness: Diarrhea that lasts 2 weeks; and
- Death

The probability of becoming infected is assumed to depend on the number of oocysts consumed during a 12-week period, the assumed minimum duration between infections by this pathogen (Section 5.5.1.). Sections 5.5.2. through 5.5.4. quantify conditional probability of becoming mildly ill, moderately to severely ill, or dying as a result of *Cryptosporidium* infection. Specifically,

- The probability of mild illness in a 12-week period is the probability of infection during that period multiplied by the conditional probability of mild illness given infection;
- The probability of moderate to severe illness in a 12-week period is the probability of mild illness during this period multiplied by the conditional probability of moderate to severe illness given mild illness; and
- The probability of death during a 12-week period is the probability of moderate to severe illness during this period multiplied by the conditional probability of death given moderate to severe illness.

The infection dose-response function and the values of conditional probabilities discussed in Sections 5.5.2. through 5.5.4. depend on AIDS status.

5.5.1. Probability of Infection. Haas et al. (1996) estimated the probability of infection as a function of oocyst intake using the exponential dose-response function. This model specifies that the probability of infection is $1 - e^{-rN}$, where N is the number of oocysts ingested, and r is the dose response function infectivity parameter.

The correct calculation of the probability of infection must take into account the fact that infection by an organism cannot occur more than once during a sufficiently short period of time. Referring to waterborne pathogens, Hurst et al. (1996) note (p. 117) that, "*It is possible for reinfection to occur as soon as 12-weeks after initial infection.*" It is therefore assumed that it is not possible to become infected more than once in a 12-week period. Hence, the probability of infection depends on the number of organisms consumed in a 12-week period. The number of organisms consumed during a 12-week period is the product of:

- The number of days in a 12-week period (84);
- The consumption rate (L/day) of unheated tap water (Section 5.2.); and
- The *Cryptosporidium* oocyst concentration in tap water (organisms / L).

Perz et al. (1998) state that for the general population, the central estimate for r is 0.0042 and that the 95% confidence interval for this parameter ranges from 0.0017 to 0.0105. This interval is symmetric about the central estimate in log space, indicating that its distribution can be characterized as lognormal with a geometric mean of 0.0042. Given the confidence interval specified, the geometric standard deviation equals 1.59[5]. For the AIDS subpopulation, Perz et al. (1998) estimate that the infectivity parameter is a factor of 3 greater than it is for the general population. For this subpopulation, it is therefore assumed that the infectivity parameter is

[5] The geometric mean of 0.0042 divided by $1.59^{1.96} = 0.0017$; likewise, the product of the 0.0042 and $1.59^{1.96} = 0.0105$.

lognormal with a geometric mean of 3 x 0.0042 = 0.0126 and a geometric standard deviation of 1.59.

5.5.2. Conditional Probability of Mild Illness Given Infection. Table 2 in Perz et al. (1998) reports the conditional probability of diarrheal illness given *Cryptosporidium* infection and the conditional probability of moderate-to-severe illness given diarrheal illness. They state that for the general population, the probability of diarrheal illness given infection is 0.40 (95% confidence interval of 0.20 to 0.80). This interval is symmetric about the central estimate in log space, indicating that it can be characterized by a lognormal distribution with a geometric mean of 0.40 and a geometric standard deviation of approximately 1.41 (the approximate square root of 2)[6].

For the AIDS subpopulation, Perz et al. report the probability of diarrheal illness given infection to be 0.95 (95% confidence interval of 0.80 to 1.00). Since the central estimate is not at the geometric center of the 95% confidence interval, it is assumed that this parameter follows a triangular distribution with bounds at 0.80 and 1.0, and a mode at 0.95[7].

5.5.3. Conditional Probability of Moderate to Severe Illness Given Mild Illness. For the general population, Perz et al. (1998) estimate that the probability of a moderate to severe illness given that the person has a mild illness has a central estimate value of 0.15 and a 95% confidence interval of 0.08 to 0.30. Because the central estimate is at the approximate geometric mean of the

[6] Since the lognormal distribution extends beyond 1.0 (the maximum possible value for this parameter), it is truncated at 1.0 and renormalized so that the area under the curve between 0 and 1.0 is unity. This adjustment introduces little distortion since a randomly drawn value from a lognormal distribution with a geometric mean of 0.40 and a geometric standard deviation of 1.41 exceeds 1.0 with a probability of only 0.4%. Alternatively, a beta distribution could have been used to characterize the range of plausible values. However, it is not clear how to specify that distribution's parameters with the information provided by Perz *et al.* (1998) – *i.e.*, the parameter's central estimate and its 95% confidence interval.

[7] The triangular distribution thus specified is not optimal since it assigns no mass to values outside the 95% confidence interval. However, given the information provided by Perz et al., there is no clear superior alternative to this specification.

95% confidence interval, the case study models this quantity as lognormal with a geometric mean of 0.15 and a geometric standard deviation of 1.41 (the approximate square root of 2)[8].

For the AIDS subpopulation, Perz et al. report estimate the corresponding probability has a central estimate value of 0.95 and a 95% confidence interval from 0.80 to 1.00. Since the central estimate is not at the geometric mean of the 95% confidence interval, a lognormal distribution is not appropriate. The case study therefore characterizes this parameter as having a triangular distribution with bounds of 0.80 and 1.0, and a mode of 0.95.

5.5.4. Conditional Probability of Death Given Moderate to Severe Illness. Eisenberg et al. (1998) report that there were 403,000 reported cases of watery diarrhea in the greater Milwaukee area during the 1993 cryptosporidiosis outbreak. Of those cases, there were 46 deaths among members of the AIDS subpopulation, and 8 deaths among members of the general population. For either the general population or the AIDS subpopulation, the conditional probability of death given moderate to severe illness can be estimated as the number of reported deaths in the population divided by the number of individuals in the population who developed a moderate to severe case of illness. However, inferring the probability of death given moderate to severe illness from this fraction is uncertain for two important reasons. First, the incidence of moderate to severe illness (the denominator of the fraction) is uncertain. Second, the fraction represents only a sample estimate of the probability. The following discussion addresses each of these sources of uncertainty.

5.5.4.1 Number of Individuals with Moderate to Severe Illness — General population: As noted above, Eisenberg et al. (1998) reports that there were 403,000 cases of

[8] Again, the lognormal is truncated at 1.0, although the probability that a quantity with a geometric mean of 0.15 and a geometric standard deviation exceeds 1.0 is approximately 2×10^{-8}.

watery diarrhea during the 1993 Milwaukee outbreak. It is plausible that many other individuals became ill but did not seek medical assistance. The Perz et al. (1998) model of cryptosporidiosis stipulates that individuals do not seek medical assistance until their illness becomes moderate to severe. It is therefore assumed that 403,000 members of the general population developed moderate to severe illness.

AIDS subpopulation: U.S. EPA is unaware of data quantifying how many members of the AIDS subpopulation developed a moderate to severe case of cryptosporidiosis. However, the conditional probability estimates from Perz et al., discussed in Sections 5.4.2. and 5.4.3., indicate that members of the AIDS subpopulation may be 15 times as likely to develop moderate to severe illness given infection than are members of the general population[9]. Since approximately 1 in 4 members of the general population developed a moderate to severe case of this illness, 403,000 of the approximately 1.64 million in the Milwaukee-Racine area (CMSA, U.S. Bureau of the Census, 1997, Table 43), it is likely that the vast majority of the AIDS subpopulation in Milwaukee developed a moderate to severe case of the illness.

Although U.S. EPA is unaware of statistics specific to metropolitan Milwaukee for that time period, Table 23 in CDC (1994a) reports that 862 individuals were living with AIDS in the state of Wisconsin in December of 1993. It is reasonable to assume that the vast majority of these individuals lived in metropolitan Milwaukee. For example, Table 2 in CDC (1997) reports that, as of December, 1997, only 8% of all reported AIDS cases in the United States were from either

[9] Perz et al.(1998) report a general population central estimate for the conditional probability of mild illness given infection of 0.40 and a general population central estimate of the probability of moderate to severe illness given mild illness of 0.15, the product of which is 0.06. The corresponding probabilities for the AIDS subpopulation are both 0.95, the product of which is 0.90.

outlying counties of metropolitan areas (2.2%) or non-metropolitan areas (5.7%).[10] The remainder were from the central counties of metropolitan areas with either 50,000 to 500,000 individuals (9.9%) or the central counties of metropolitan areas with more than 500,000 individuals (82.2%). In summary, it is assumed that all 862 members of the AIDS subpopulation developed moderate to severe cryptosporidiosis during the 1993 Milwaukee outbreak.

5.5.4.2. Sampling Uncertainty — Berger (1980, p. 287) states that the binomial distribution parameter, which in this context is the conditional probability of death given moderate to severe illness, can be described as following a beta distribution[11]. Assuming there are n deaths out of a population of size p, the beta distribution's parameters are $\alpha = 1+n$, and $\beta = 1+(P-n)$[12]. For the general population, the beta distribution describing the relative likelihood of plausible values for the conditional probability of death has parameters $\alpha = 9$ (i.e., 1 + 8), and $\beta = 402{,}993$ (i.e., 1+ 403,000 - 8). The mean of this distribution which is illustrated in Figure 5-5 is $9 \div 403{,}002$, or approximately 2.2×10^{-5}. The corresponding distribution for the AIDS subpopulation appears in Figure 5-6 and has parameters $\alpha = 47$ (i.e., 1 + 46), and $\beta = 817$ (i.e., 1 + 962 - 46). Its mean is $47 \div 864$, or 5.4%.

[10]The reporting of cases of acquired immunodeficiency syndrome (AIDS) is mandatory. All 50 states and the District of Columbia report cases to the CDC using a uniform case definition and reporting form. The definition of an AIDS case has changed over time and therefore rate comparisons over time must be interpreted very cautiously. Recently available treatment regimens have resulted in increased survival of AIDS patients with a concomitant increase in the prevalence of AIDS cases. The CDC has maintained and published statistical information on the AIDS epidemic since its beginning. The data in the current report on prevalence of AIDS in specific geographic areas was abstracted from tables reported in various versions of the HIV/AIDS Surveillance Report, now published biennially by the CDC.

[11] The beta distribution family assigns weight only to values between 0 and 1 – the range of plausible values for the binomial distribution parameter estimate. It can take on a variety of shapes.

[12] Technically, Berger states that if the prior distribution for a parameter follows a beta distribution with parameters a and b, then the posterior distribution reflecting both the prior and the sample information follows a beta with a distribution with $\alpha = a + n$, and $\beta = b + (P - n)$. A prior that assigns equal likelihood to all potential binomial parameter values between 0 and 1 is the uniform distribution. The uniform distribution is a special case of the beta distribution with $\alpha = 1$, and $\beta = 1$.

Figure 5-5
Probability Density Function for the Conditional Probability of Death Given Moderate to Severe Illness: General Population

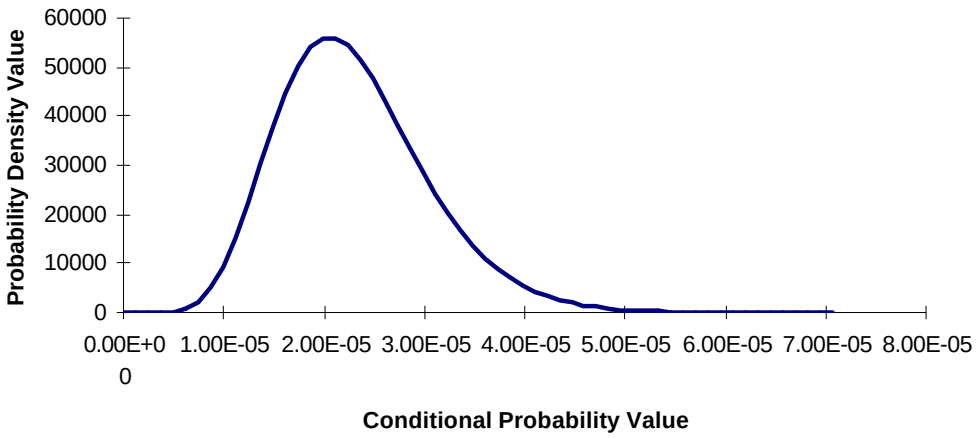
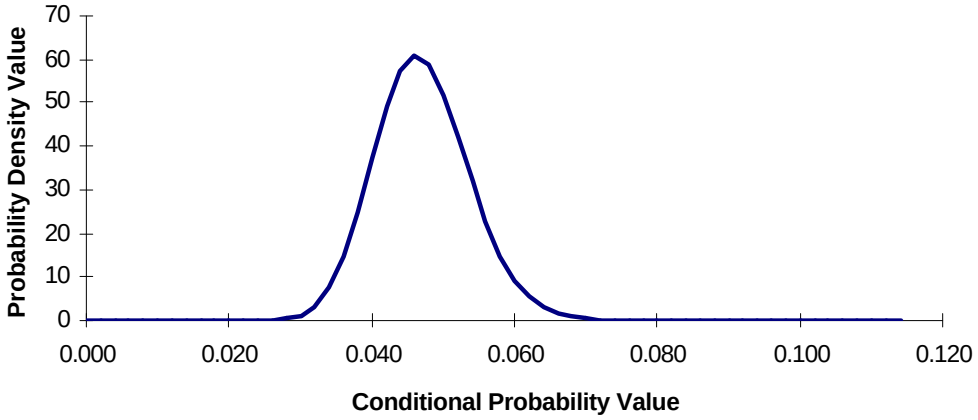


Figure 5-6
Probability Density Function for the Conditional Probability of Death Given Moderate to Severe Illness: AIDS Subpopulation



5.6. TREATMENT TECHNOLOGY COSTS

Table 5-18 illustrates the financial costs of the baseline and supplemental treatment technologies considered in the case study. These costs are limited to estimated initial implementation and annual maintenance and operating costs. Appendix A.1 discusses the basis for the estimates. Table A.1-3 contains the assumptions underlying the cost analysis. For a 130 MGD treatment plant, the incremental capital cost for ozone pretreatment is $4,800,000. Annual operation costs are $400,000. For a population of 460,000 individuals, the per capita capital cost is $10.43, while the per capita operational costs are $0.87. Lykins et al. (1991) estimated the costs of an in-home reverse osmosis filter system to be $500 to $800 capital expenditure and an annual maintenance cost of $500 to $150 (in 1991 dollars). Equivalent costs in 1996 dollars (the most recent year for which information is available) are $575 to $920 for capital costs and $80.50 to $172.50 for annual operating costs (U.S. Bureau of the Census, 1997, Table 751). For purposes of this case study, costs are assumed to equal the midpoint of the ranges, or approximately $750 and $125, respectively.

5.7. DISCOUNT RATE

In the current study, a social discount rate of 3% is used for converting future costs and health benefits into present value equivalents. This rate is consistent with the range of 2 to 3% for estimates of both the long term rate of return on investment (Hartman, 1990; U.S. OMB, 1996b), and the social time preference for consumption (Moore and Viscusi, 1990; Lind, 1990; Freeman, 1993). Three percent was chosen as a point estimate to enhance comparability of the reported results to other studies. Benefits and costs were also calculated using a discount rate of 5% in the sensitivity analysis to facilitate comparison with other studies.

Table 5-18

Treatment Technology Costs

TREATMENT	INITIAL COST	ANNUAL COST
BASELINE (Chlorination)[a]	$78,200,000	$7,300,000
BASELINE WITH PRE-OZONATION[a]	$4,800,000	$400,000
IN-HOME REVERSE OSMOSIS FILTER	Purchase: $500 - $800 Installation $70 - $150	$50 - $100

Notes: Initial estimated construction costs (not amortized) are given in 1997 dollars without adding interest .

Lykins et al., (1992 p. 205) estimated the costs (in 1991 dollars) of a reverse osmosis unit for a single tap.

Gold et al. (1996) recommend a discount rate of 3% for discounting both costs and health benefits in cost-effectiveness studies of public health interventions. Haddix et al. (1996) suggest that a rate of 3 to 5% would be appropriate. Both note that reporting results based on a range of discount rates is recommended, in order to insure comparability between studies.

5.8. THE COST OF HEALTH EVENTS IN QUALITY ADJUSTED LIFE YEARS (QALYS)

This section estimates the cost of the health effects evaluated in the case study in terms of lost Quality Adjusted Life Years (lost QALYs). Briefly, the factors affecting the magnitude of this cost include the following:

- The severity of the health effect and its duration;
- The number of lost life years associated with the health effect; and
- The rate at which future health benefits (and costs) are discounted.

Section 5.8.1. describes data used to quantify QALY equivalent values for various health outcomes considered in this case study. Section 5.8.2. applies this information to derive specific

QALY values (or expressions for these values that depend on other parameters) for cancer, reproductive toxicity, developmental toxicity, illness associated with cryptosporidiosis, and mortality associated with cryptosporidiosis.

5.8.1. Data Used to Derive QALY Equivalent Costs. This section reviews literature that can be used to estimate the lost QALYs associated with various adverse health states. The following discussion first reviews results reported by the Beaver Dam Health Outcome Study (Fryback et al., 1993), also referred to as the "BDHOS" (Section 5.8.1.1.). Also reviewed are results from a survey conducted among individuals living in Hamilton, Ontario (Torrance et al., 1992) (Section 5.8.1.2.). Finally, a collection of studies are discussed that develop estimates for the QALY cost of cancer (Section 5.8.1.3.).

5.8.1.1. The Beaver Dam Health Outcome Study (BDHOS) — The BDHOS reports results for 1,356 non-institutionalized participants drawn from a cohort of 43-to 84 year-old individuals living in Beaver Dam, Wisconsin. Each participant specified the amount of time in the ideal state of perfect health that was equally desirable to living with his or her current health conditions for that individual's remaining life expectancy. The ratio of this duration in perfect health to that individual's actual life expectancy equals the QALY value that individual places on each year remaining in his or her life. For example, if an individual is indifferent between 15 years of life in perfect health and 20 years of life with his or her health conditions, he or she is indifferent between each year of life in his or her present health state and $15 \div 20 = 0.75$ years in perfect health. That is, each year of life has a value of 0.75 QALYs, meaning that the "cost" of his or her health conditions is 0.25 QALYs per year.

The BDHOS reports the QALY value of each year of life for individuals with each of 28 conditions. It must be noted that it would be incorrect to estimate the QALY cost of each of these listed health conditions by subtracting from 1 the QALY value of a single life year for individuals suffering from that condition. This difference represents the cost of all the adverse health conditions suffered by the individual, rather than the cost of the one health condition under consideration. For example, individuals with arthritis place a value of 0.815 QALYs on each year of life, meaning that, due to imperfect health, they lose 0.185 QALYs each year. However, this loss reflects both the impact of arthritis and the impact of other typical conditions for this population. That is, the cost of arthritis is some value less than 1 - 0.815 = 0.185 QALYs per life year.

The ideal estimate of this quantity is the difference between the QALY value placed on a year of life by individuals with the health condition under consideration (referred to here as the "sub-population" suffering from a specified health condition) and the QALY value individuals from an average "reference" population place on each year of life. The reference population is a hypothetical population of individuals who have the same set of health disorders as the sub-population with the exception of the condition whose value is to be estimated. The assumption underlying use of the reference population is that, with the exception of the health condition of interest, their health is, on average, equally desirable to the health of the sub-population suffering from this condition. In short, the sub-population suffers from the "average" set of health conditions plus the health condition under consideration, while the reference population suffers only from the average set of health conditions. Hence, the difference between the QALY value

assigned to a year of life by the reference population and the QALY value assigned to a year of life by the sub-population reflects the cost of the health condition under consideration.[13]

Since the value placed on a year of life by this hypothetical reference population is not known, a series of proxies using the BDHOS study sample have been created. Specifically, for each health condition, the "reference" population consists of those subjects in the BDHOS sample who do not suffer from that condition. The cost of each health condition therefore equals the difference between the value of a year of life for individuals with each of these conditions and the value of a year of life for individuals without each of these conditions. Table 5-19 details these calculations. The cost of angina, for example, was calculated as follows. A total of 65 of the BDHOS participants reported that they suffered from angina, along with other conditions. The average value of a year of life for these individuals was 0.786 QALYs. A total of 1,253 participants reported that they did not suffer from angina. For this "reference population," the average value of a year of life was 0.864 QALYs. The annual cost of angina is estimated to be the value of a year of life among those who do not suffer this disease (0.864 QALYs) minus the value of a year of life among those who do suffer from this disease (0.786), or 0.078 QALYs. This value appears in the far right column of Table 5-19.

Use of a proxy reference population, as described in the preceding paragraph, complicates the interpretation of the analysis. First, a different reference population is used to evaluate the cost of each health condition. That is, there is no constant benchmark to which each sub-

[13] Alternatively, one could regress (using ordinary least squares regression) the QALY value individuals place on each year of life against a set of dichotomous variables, each of which indicates whether an individual has a specified health condition. The magnitude of the coefficient for each indicator variable would equal the QALY cost of the corresponding health condition. The published BDHOS results do not provide the information necessary to conduct this type of analysis.

Table 5-19
Fraction of Life-span in Perfect Health Equally Desirable to Full Life-span with Current Health Status[a]

Condition	Participants with Condition		Participants without Condition		Lost QALYs per Year of Life in this Condition
	n	Value of one year of life in QALYs	n	Value of one year of life in QALYs	
Arthritis	598	0.815	720	0.900	0.085
Gout	56	0.859	1262	0.861	0.002
Severe back pain	240	0.786	1078	0.878	0.092
Severe neck pain	99	0.765	1219	0.869	0.104
Migraine	73	0.817	1245	0.864	0.047
Angina	65	0.786	1253	0.865	0.079
Congestive Heart Failure	28	0.710	1290	0.865	0.155
Myocardial infarction	20	0.729	1298	0.863	0.134
Stroke	10	0.903	1308	0.861	NA[b]
Hypertension	479	0.830	839	0.879	0.049
Hyperlipidemia	109	0.902	1209	0.858	NA[b]
Cataract	314	0.821	1004	0.874	0.053
Glaucoma	66	0.824	1252	0.863	0.039
Macular degeneration	38	0.754	1280	0.864	0.110
Diabetes (insulin)	35	0.627	1283	0.868	0.241
Diabetes (no insulin)	82	0.761	1236	0.868	0.107
Asthma	46	0.706	1272	0.867	0.161
Emphysema	38	0.751	1280	0.865	0.114
Chronic bronchitis	46	0.724	1272	0.866	0.142
Chronic sinusitis	92	0.874	1226	0.860	NA[b]
Depression	60	0.703	1258	0.869	0.166
Anxiety	52	0.774	1266	0.865	0.091
Ulcer	75	0.790	1243	0.866	0.076
Colitis	51	0.815	1267	0.863	0.048
Hiatal hernia	44	0.845	1274	0.863	0.018
Sleep disorder	135	0.790	1183	0.869	0.079
Thyroid disorder	82	0.882	1236	0.860	NA[b]
Misc. Allergies	28	0.844	1290	0.861	0.017

Source: Fryback et al. (1993), Table 4

Notes: NA" indicates that the estimate of lost QALYs per life year is negative and hence not valid. See text.

population suffering from a different disease is compared. Even more importantly, the proxy reference population does not, as ideally required, have all the health conditions of the study population, with the exception of the health condition under consideration. That is, members of the reference population may suffer from additional health conditions not suffered by the study population. These additional health conditions may be sufficiently severe to yield a life year value for the reference population that is lower than the life year value for the sub-population. In these cases, the subtraction described in the previous paragraph yields a negative value. For example, individuals in this sample who had a thyroid condition placed a value of 0.882 QALYs on each year of life, while those without this condition placed a value of 0.860 QALYs on each year of life (see Table 5-19). For the reasons just stated, this result does not mean that having a thyroid condition improves an individual's quality of life. In these cases, the entry in the right column of Table 5-19 is marked with "NA".

It must also be noted that this bias can operate in the opposite direction. Specifically, the health conditions suffered by the reference population may, in aggregate, be less severe (as measured in QALYs) than the health conditions that would be suffered by the sub-population in the absence of the health condition under consideration. In this case, the difference between the value of each life year for the reference population and the value of each life year for the sub-population overstates the cost of the health condition under consideration.

In general, the calculations in Table 5-19 overstate the cost of some conditions, and underestimate the cost of others. With the exception of the costs that have been calculated to be negative, it is not possible determine, from the information available from the BDHOS data, for

which health conditions the cost estimate has been biased either positively or negatively. In no case is it possible to estimate the magnitude of this bias.

5.8.1.2. The Hamilton, Ontario Study — Torrance et al. (1992) identified 718 children from kindergarten through grade 5 attending publicly funded schools in the city of Hamilton, Ontario. The investigators interviewed a parent for each of a subset of these children, resulting in 293 sets of responses. Of these, 203 were judged to be useful. The other 90 sets of responses were omitted from further consideration because either certain key responses were missing, the interviewer judged the interview to be of poor quality, or the responses did not satisfy certain, predefined criteria of internal logical consistency.

The survey asked the 203 Hamilton Ontario respondents to rate the relative desirability of poor health (the nature of which was described to the survey participants) for each of seven attributes: sensation (including the ability to see, hear, and speak normally); mobility, emotion, cognition, self-care, pain, and, fertility. Each participant assigned a rating on a visual scale ranging from 0 (corresponding to death) to 1 (corresponding to perfect health) assuming that he or she suffered from the worst state of health for that attribute, but enjoyed perfect health for the other 6 attributes. By assessing the value corresponding to each adverse health state, along with the utility of a limited combination of adverse health states, Torrance et al. (1992) established a relationship between the health state values directly surveyed and the lost utility associated with each adverse health state. Specifically, they report that the health state values elicited in this

study raised to the power 2.29 can be used to approximate utility (which are measured here in units of QALYs)[14]. For our purposes, three findings are useful:

- Fertility: The inability to have children with a fertile spouse was assigned an average utility cost of 0.12. This result means that individuals were on average indifferent between living a full year in this infertile state, and living in perfect health for one year with a probability of 88% and losing one year of life. That is, these findings indicate that each year of infertility is equivalent to a loss of 0.12 QALYs.

- Pain: Severe pain (pain not relieved by drugs and constantly disruptive of normal activities) had a utility cost of 0.36. That is, a year in severe pain has a value of 0.64 QALYs.

- Self Care and Mobility: Extreme dependency (requiring the help of another person to eat, bathe, dress, or use the toilet), had a utility cost of 0.45. That is, a year in this condition has a value of 0.55 QALYs.

5.8.1.3. Studies Providing QALY Cost Estimates for Cancer — This section briefly reviews a collection of studies that have developed estimates of the cost associated with cancer (i.e., the value, measured in QALYs, of a year of life with cancer), typically in the context of conducting a cost effectiveness analysis.

Norum et al. (1997) evaluated the cost-effectiveness of adjuvant chemotherapy as part of the treatment for colorectal carcinoma. Investigators sent surveys to 95 cancer patients, of whom 62 responded. The patients were asked to rate their quality of life using three instruments. The EuroQol questionnaire asks subjects to evaluate their health along 5 dimensions (mobility, self-care, daily activities, pain, and mood), rating each as "no problem," a "moderate problem," or an

[14] To the knowledge of U.S. EPA, this relationship has not been independently verified by other researchers, a factor that introduces uncertainty into the QALY estimates derived from this study's results. However, it is also true that QALY estimates derived by other investigators also suffer from problems that make them uncertain. Problems include: inability of study subjects to express their true preferences in a hypothetical setting; inability of subjects to comprehend the preference elicitation questions; and use of proxies (*e.g.*, doctors or other experts) in the place of patients suffering from some health effect.

"extreme problem." In a separate analysis conducted by the EuroQol group (Williams, 1990), investigators used a visual analogue scale to assign scores between 0 and 1 to each combination of health states[15]. Norum et al. (1997) also assessed the value of a year (in QALYs) with cancer by directly using a visual analogue scale. Finally, Norum et al. (1997) used the global QoL-measure developed by the European Organization for Research and Treatment of Cancer, a scale that asks respondents to rate their health on a 1 to 7 scale (Aaronson et al., 1993). It should be noted that none of these scales are true utility measures since respondents are not asked to consider a trade-off of some benefit (e.g., extended life) in exchange for elimination of an adverse health condition. In any case, all three instruments yielded the same result for the study subjects – *i.e.*, each year with cancer had a value of 0.83 QALYs.

A study conducted by Bennett et al. (1996) suggests that the value of a year with cancer depends, not surprisingly, on the severity of the symptoms. Bennett et al. report the time trade-off judgments expressed by four focus groups of urologists asked to assess the value of a year of life for various stages of metastatic prostate carcinoma. For these stages, the median value of a single year of life were: 0.92 QALYs for stable disease, 0.84 QALYs for stable disease with gastrointestinal toxicity due to treatment, 0.83 QALYs for early progressive disease, and 0.42 QALYs for late progressive disease. The validity of these values is somewhat compromised by the use of physicians as proxies for patients.

Norum et al. (1996) used the EuroQol scale (see discussion of Norum et al. (1997), above) to assess the QALY value of a year of life for four different stages of Hodgkins Disease.

[15] The visual analogue scale is a visual aid, like a thermometer, with 0 at one end, representing death, and 1 at the other end, representing perfect health. The respondent identifies the point between these two extremes that represents his or her perception of the specified health state.

The results ranged from 0.73 to 0.81 QALYs per life year (i.e., cancer "costs" ranged between 0.19 and 0.27 QALYs per year), although the value of a life year did not monotonically decrease with disease stage, as one would expect. The average value of a year with cancer was 0.78 QALYs. It should be noted that in this case, the authors state that the time trade-off methodology was used to assess the QALY value of each of the combination of health states evaluated as part of the EuroQol scale. The methodology section does not clarify the inconsistency between this description of the EuroQol survey and the description provided by Norum et al. (1997), which states that a visual analogue scale was used to estimate QALY values for various health states.

Grunberg et al. (1996) surveyed 30 cancer patients to determine the cost, in terms of QALYs, of nausea and vomiting associated with chemotherapy. The assessment was conducted using a visual analogue scale, which for theoretical reasons mentioned earlier, does not constitute a valid measure of utility in terms of QALYs. In any case, Grunberg et al. found that severe nausea and vomiting can substantially depress the value of life. Patients judged that in the absence of these symptoms, the value of one year of life was 0.79 QALYs; with nausea and vomiting, this value dropped to 0.27 QALYs.

Hutton et al. (1996) used nurses as proxies for breast cancer patients. Using the standard gamble elicitation approach (Raiffa, 1968), they assessed the value of a year of life with this cancer given different treatment outcomes. The value of each life year ranged from 0.81 QALYs (partial response to treatment) to as little as 0.13 QALYs (terminal disease).

5.8.2. QALY Equivalent Values Derived. This section uses the health preference data described in Section 5.8.1 to estimate lost QALYs for each of five health conditions: death due to

either cancer or microbial infection (Section 5.8.2.1.), cancer illness (Section 5.8.2.2.), developmental toxicity (Section 5.8.2.3.), reproductive toxicity (Section 5.8.2.4.), and gastrointestinal illness caused by *Cryptosporidium* infection (Section 5.8.2.5.). Section 5.8.2.6. summarizes these QALY cost assignments, and describes how the case study characterizes uncertainty for these estimates.

Where appropriate, adjusted estimates are derived for the AIDS subpopulation. To this end, two simplifying assumptions are made. First, it is assumed that the value of a year of life for members of the AIDS subpopulation is 0.6 QALYs. U.S. EPA is unaware of data that specifically addresses this parameter. However, the proposed value is similar to the value of 0.55 QALYs reported by the Hamilton, Ontario study (Torrance et al., 1992) for conditions resulting in extreme dependency. It is also similar to the value of 0.64 QALYs reported by Torrance et al. for conditions producing severe pain. Second, it is assumed that members of the AIDS subpopulation (as distinct from the HIV-positive subpopulation, members of whom carry the HIV virus but do not manifest the AIDS disease) do not bear children. Hence, only cancer and pathogen-induced illness and mortality are relevant to this subpopulation.

5.8.2.1. Death — The QALY cost of death is the net present value of the years of life that are lost. Equation 5-8-1 quantifies the net present value of the QALY cost of death at age D ($Val(Death)_D$) in the year that it occurs as:

$$Val(Death)_D = \sum_{i=D}^{MaxAge} Q_i \times \frac{1}{(1+d)^{i-d}} \times PrLive_{D,i} \qquad (5\text{-}2)$$

where:

D	=	Age of death
MaxAge	=	Maximum age to which members of the population might live (assumed to be 89 for the purpose of this study)[16]
Q_i	=	QALY value of a year of life at age i
d	=	The annual discount rate.
$PrLive_{D,i}$	=	The conditional probability that an individual who lives to age D will live to at least age i.

Note that, as described in Section 6.2., there is a latency period between exposure to a carcinogen and the onset of the resulting cancer. To calculate the net present value of the QALY cost of death due to cancer at the time when exposure occurs (as opposed to its net present value at the time of death), the preceding expression must be discounted by an amount corresponding to this latency period. In other words, a death occurring at some age has a greater cost if it occurs immediately than if it is delayed for some period of time. For example, it is worse to contemplate death at age 75 when one is 70 than when one is 40.

Q_i: Table 5-20 summarizes the assumed QALY values for life years at different ages (*i.e.*, it describes the function, Q_i for members of the general population and AIDS subpopulation).

$PrLive_{D,i}$: Table 5-21 summarizes the probability that an individual alive at the beginning of the age range in the left column will be alive at the end of that age range. For example, among the general population, an individual who has just turned age of 55 will be alive

[16] A value of 89 has been selected because few statistics are reported for higher age groups. Moreover, increasing the value of MaxAge beyond 89 would not substantially alter the analysis since so few individuals live beyond this age.

Table 5-20

The QALY Value of Life Years at Different Ages

Age Range (years)	QALY Value of Each Life Year	
	AIDS Subpopulation	General Population
0 to 14	0.6[a]	0.950[b]
15 to 24	0.6[a]	0.950[b]
25 to 34	0.6[a]	0.950[b]
35 to 44	0.6[a]	0.950[b]
45 to 54	0.6[a]	0.921[c]
55 to 64	0.6[a]	0.873[c]
65 to 74	0.6[a]	0.837[c]
75 to 84	0.6[a]	0.815[c]
85	0.6[a]	0.808[c]

Notes: [a]*Assumed value.*

[b]*Life years before age 45 are assumed to be equivalent to 0.95 QALYs. Fryback et al. Did not report values for these age groups.*

[c]*Reported in Fryback et al. (1993), Table 2.*

with probability 0.957 at age 60. For the AIDS subpopulation, the probability of surviving any 5-year period is assumed to be 0.531.

The survival probabilities for the general population have been taken directly from the life table statistics published by NCHS (1998).

Several factors complicate the development of survival probability for the AIDS subpopulation. First, the distribution of duration since diagnosis (and hence the distribution of disease severity) changes over time. Second, medical advances continue to lengthen life expectancy for the AIDS subpopulation. However, as of 1996, the annual fatality rate among

Table 5-21

Probability of Surviving Until the End of Each Five-Year Age Range for Individuals Alive at the Beginning of that Age Range

Age Range	General Population[a]	AIDS Subpopulation[b]
0 to 4	0.991	0.531
5 to 9	0.999	0.531
10 to 14	0.999	0.531
15 to 19	0.996	0.531
20 to 24	0.995	0.531
25 to 29	0.994	0.531
30 to 34	0.992	0.531
35 to 39	0.990	0.531
40 to 44	0.986	0.531
45 to 49	0.981	0.531
50 to 54	0.972	0.531
55 to 59	0.957	0.531
60 to 64	0.933	0.531
65 to 69	0.902	0.531
70 to 74	0.854	0.531
75 to 79	0.787	0.531
80 to 84	0.681	0.531
85+	0.000	0.531

Notes: [a]*Based on Page 5 of Section 6 in NCHS (1998).*

[b]*See text.*

members of the AIDS subpopulation was approximately 11.9%. Specifically, there were approximately 247,000 individuals living with AIDS in 1997 (Table 33 in U.S. CDC, 1997). The 1996 AIDS population size is likely to be similar. In 1996, the fatality rate due to AIDS for the entire population of the United States was 11.1 per 100,000 (Table 44 in Pamuk et al., 1998). The population of the United States that year was approximately 265,000,000 (Table 14 in U.S. Bureau of the Census, 1997). Hence, approximately 93 of every 100,000 individuals in the population had AIDS. Since the AIDS death rate for the entire population was 11.1 per 100,000,

the death rate per member of the AIDS subpopulation was approximately 11.1 ÷ 93, or 11.9%. This value translates into a 5-year survival probability of 0.531, which is assumed to apply across all age groups.

Tables 5-22 (general population) and 5-23 (AIDS subpopulation) detail the conditional probability of surviving until at least the age range listed in the far left column of each row given that an individual is alive at the age listed at the top of each column. For example, an individual between the ages of 5 and 9 has an 86.8% probability of surviving until at least age 60 to 64.

5.8.2.2. Cancer Illness — The cost of nonfatal cancer illness expressed in QALYs depends on the duration of the illness and the severity of the symptoms and treatment. For the purpose of this study, it is assumed the duration of the illness is 2 years.

Illness severity: An appropriate QALY cost per year of cancer illness can be estimated on the basis of the studies described in Section 5.8.1. Although 60 of the 1,356 participants in the BDHOS reported that they had been affected by cancer in the preceding year, results for cancer are not reported by Fryback et al. (1993). The authors state that of the 60 cancers, 31 were skin cancers (excluding melanoma), and another 8 were cancer of the breast. However, the authors do not state why results for this disease are not reported.

To estimate the cost of cancer using the BHDOS data, an attempt has been made to dentify an alternative, severe, chronic condition to serve as a proxy for cancer. As cancer is typically a severe condition with a number of adverse symptoms, an appropriate proxy must be among those conditions associated with the greatest QALY decrements. Among those conditions with the largest QALY costs per year of life are: arthritis (0.085), severe back pain (0.092), severe neck pain (0.104), macular degeneration (0.110), diabetes with insulin (0.241),

diabetes without insulin (0.107), asthma (0.161), emphysema (0.114), chronic bronchitis (0.142), and depression (0.166). These costs tend to range from approximately 0.10 QALYs per year to as much as approximately 0.25 QALYs per year. Assuming that treatment of active cancer lasts for 2 years, the cost reflecting the resulting decreased quality of life associated with cancer is between 0.2 and 0.5 QALYs.

The QALY cost of cancer morbidity can also be inferred from the results of the Hamilton, Ontario study (Torrance et al., 1992). As cancer itself, along with various treatments for cancer, such as chemotherapy, can involve extreme pain, we used the severe pain cost of 0.36 QALYs per year reported by Torrance et al.

The cost effectiveness studies described in Section 5.6.1.3. indicate that the symptoms associated with cancer and its treatment can decrease the value of a year of life to as little as 0.13 QALYs (see Hutton et al., 1996). In other cases, the value of a year of life could be relatively high – e.g., 0.83 QALYs (see Norum et al., 1997). For the purpose of this assessment, it will be assumed that relatively aggressive cancer treatment proceeds for 2 years, after which time, the patient either dies or the cancer goes into remission. Corresponding to the assumption of relatively aggressive cancer therapy during this period, it will be assumed that the value of each year of life during treatment is 0.5 QALYs. Hence, the cost of cancer is 2 years × (QALY value of a life year in typical health - QALY value of life during treatment).

For the general population, the first term in the parentheses is age-dependent. In this case, the cost of the decreased quality of life associated with cancer is estimated to be 2 years × (Q_i - 0.5) QALYs, where Q_i is the value of a year of life at age i (see Table 5-19). Since the value of a year of life for the AIDS subpopulation is assumed to be 0.6 QALYs (see Table 5-19), the cost of

the decreased quality of life associated with cancer is estimated to be approximately 2 × (0.6 - 0.5) QALYs, or 0.2 QALYs.

5.8.2.3. Developmental Toxicity Associated with DBP Ingestion — Calculation of the QALY value for developmental toxicity is complicated by the fact that it affects both an individual to be born and the lives of that child's parents. For the purpose of this study, only the former cost category has been included in the estimation of the QALY cost associated with developmental toxicity endpoints. Sensitivity analysis can be used to determine if consideration of the costs to the parents (which can of course be substantial) are likely to affect the analysis.

Calculation of the QALY costs for developmental toxicity is further complicated by the fact that this category encompasses a wide range of disparate health effects. For example, developmental toxicity may result in a miscarriage; alternatively, it may result in a live birth and some degree of developmental defects. Clearly, birth defects that are evident at birth represent only part of a spectrum of adverse outcomes that can appear later in life during childhood. Certain conditions such as mental retardation, cerebral palsy, blindness and hearing impairment are termed developmental disabilities and generally are manifested after infancy. Developmental disabilities are considered to be lifetime conditions and they can result in substantial costs to the affected individuals, their families, and society (Boyle et al., 1996). For the purpose of this analysis, U.S. EPA assumes that developmental toxicity represents a live birth with associated severe developmental defects. In order to quantitatively assess this outcome, it is further assumed that the child's defects result in a state of extreme dependency. The Hamilton, Ontario data (Torrance et al., 1992) indicate that each year of life in a state of extreme dependency has a QALY value of 0.55, assuming no other health problems.

U.S. EPA also assumes that individuals suffering from developmental defects have a 50-year life expectancy, considerably less than the 75.8 year life-expectancy of individuals without severe developmental defects (U.S. Bureau of the Census, 1997, Table 117). This assumption is highly uncertain, although it is not inconsistent with information available regarding the impact of severe birth defects on life expectancy. Although data are collected on the incidence and prevalence of birth defects and developmental disabilities, there is surprisingly little data available on the expectation of life at birth for individuals diagnosed with these conditions. A recent publication on the economic costs of birth defects and cerebral palsy indicates that life expectancy can be expected to vary by condition, although the condition-specific expectations are not directly stated (CDC, 1995b), and were not amenable to teasing out of the presented summary statistics. Medical costs for some conditions are computed through age 65, but for others they are truncated at much earlier ages, e.g., ages 9 and 17. In the absence of more specific data on life expectancy and shortening, for the purpose of the present analysis, U.S. EPA chooses the value of 50 years for the expected duration of life for persons with birth defects and/or severe developmental disabilities.

Finally, U.S. EPA assumes that individuals with developmental defects experience other adverse health effects, and that, on average, these other health effects are similar to those experienced by members of the general population. Hence, severe developmental defects reduce the value of each year of life from its age-specific baseline value (see Table 5-20) to 0.55 QALYs. Multiplication by $PrLive_{0,i}$ (the probability of living until at least age i at the time of birth) adjusts for the possibility that the individual may have died at various ages even in the absence of the developmental defect. Equation 5-3 quantifies lost QALYs due to severe developmental defects,

where d is the annual discount rate. It is assumed that members of the AIDS subpopulation do not bear children; hence this health effect is not applicable to that subpopulation.

$$Val(DevelTox) = \sum_{i=0}^{50}(Q_i - 0.55) \times PrLive_{0,i} \times \frac{1}{(1+d)^i} + \sum_{i=51}^{75} Q_i \times PrLive_{0,i} \times \frac{1}{(1+d)^i} \quad (5\text{-}3)$$

5.8.2.4. Reproductive Toxicity Associated with DBP Ingestion — As with developmental toxicity, calculation of QALY values for reproductive toxicity is complicated by the range of health effects encompassed by this category. The case study interprets reproductive toxicity to include outcomes affecting adult fertility.[17] Specifically, it is assumed that an affected individual is no longer able to conceive a child with a fertile spouse. The Hamilton, Ontario survey (Torrance et al., 1992) investigated the preferences associated with this health outcome and reports that, on average, the annual "cost" of complete infertility is 0.12 QALYs per year. Note that it is assumed that the cost is born by both members of the affected couple. That is, the cost of a single year of infertility among a couple attempting to conceive a child is 0.24 QALYs. It is assumed that members of the AIDS subpopulation do not bear children; hence this health effect is not applicable to that subpopulation.

5.8.2.5 Transitory Gastrointestinal Illness Associated with Cryptosporidiosis — The lost QALY cost associated with gastrointestinal infections depends on the duration of each infection and the impact of infection on quality of life. While neither the BDHOS (Fryback et al., 1993) nor Hamilton, Ontario study (Torrance et al., 1992) directly addresses GI illness, some of

[17] Data on birth and fertility rates are collected and generated by the National Center for Health Statistics based on the national vital statistics system of the United States and are summarized in many different publications. The information in the current report was abstracted from Pamuk et al., 1998. Registration of births in the U.S. is documented to be over 99% complete.

the results can be used to infer a quantitative health preference for this condition, and hence to estimate its QALY cost. The Hamilton, Ontario study (Torrance et al., 1992) indicates that a year of severe pain that "is not relieved by drugs and constantly disrupts normal activities" has a value of 0.64 QALYs. The cost of this condition is hence the difference between the value of a life year in the absence of GI infection and 0.64 QALYs. A mild case of GI illness is assumed to last one day, and hence has a cost equal to the value in the third column of Table 5-24 divided by 365 (see column 4 in Table 5-24). The cost of a moderate to severe case of GI illness, which is assumed to last 2 weeks, appears in the far right column in Table 5-24.

For members of the AIDS subpopulation, estimating this quantity is complicated by the fact that for these individuals, the typical value of a year of life — even in the absence of GI illness — is relatively low; for this case study, it is assumed to be 0.6 QALYs, which is less than the 0.64 QALY value of a year of life with GI infection described in the preceding paragraph. This apparent paradox of course indicates that the 0.64 QALY value for a year of life with GI illness is not applicable to the AIDS subpopulation. It is difficult to estimate how much GI illness decreases the value of a year of life for an individual with AIDS since this group already suffers from a debilitating condition. If the marginal cost of successive adverse conditions decrease, it is reasonable to assume that the cost for this condition must be no greater than the cost for the general population. In this sense, the far right column in Table 5-24 serves as an upper bound for the QALY cost for the AIDS subpopulation. On the other hand, it is unlikely that members of the AIDS subpopulation are so uncomfortable that GI illness effectively has no

Table 5-24
QALY Cost of GI Illness for the General Population

Age Group	Value of a Year of Life without GI Illness (QALYs)[a]	QALY Cost of a Year of Life With GI Illness[b]	QALY Cost of a Single Case of Mild GI Illness[c]	QALY Cost of a Single Case of Mild GI Illness[d]
0 to 14	0.950	0.31	8.49E-4	1.19E-2
15 to 24	0.950	0.31	8.49E-4	1.19E-2
25 to 34	0.950	0.31	8.49E-4	1.19E-2
35 to 44	0.950	0.31	8.49E-4	1.19E-2
45 to 54	0.921	0.281	7.70E-4	1.08E-2
55 to 64	0.873	0.233	6.38E-4	8.94E-3
65 to 74	0.837	0.197	5.40E-4	7.56E-3
75 to 84	0.815	0.175	4.79E-4	6.71E-3
85 and above	0.808	0.168	4.60E-4	6.44E-3

Notes: [a]*Values taken from Table 5-20.*

[b]*This column lists the difference between the value of a year of without GI illness and the value of a year of life with GI illness (0.64 QALYs).*

[c]*A mild case of GI illness is assumed to last 1 day; the values in this column therefore equal the values in column 3 divided by 365.*

[d]*A moderate to severe case of GI illness is assumed to last 14 days; the values in this column therefore equal the values in column 3 divided by 365.*

cost because it is "swamped" by the symptoms of AIDS. For the purpose of this case study, it is assumed that the QALY cost associated with GI illness is equal to one-half the age-weighted average cost for the general population. For mild illness, this cost is one-half the age-weighted average of column 4 in Table 5-24, or 3.60×10^{-4} QALYs. For moderate to severe illness, this cost is one-half the average of column 5 in Table 5-24, or 5.04×10^{-3} QALYs.

5.8.2.6 QALY Costs – Summary — The following lists briefly summarizes the basic approach used to assign values to each of the health endpoints considered in this case study:

- DBP-induced cancer illness: Treatment costs per year are the difference between a typical year of life and the value of a year of life undergoing cancer treatment (0.5). The treatment is assumed to last 2 years.

- DBP-induced cancer death: The QALY cost of this event is the value of the life years lost.

- DBP-induced reproductive toxicity: The cost of a single year of infertility is assumed to be 0.12 QALYs. This cost is assumed to affect both males and females. This cost is assumed not to apply to the AIDS subpopulation.

- DBP-induced developmental toxicity: Only the cost to the child to be born is considered. It is assumed that this individual will suffer a decreased quality of life due to severe dependence stemming from the developmental toxicity. The cost of this dependency is the difference between the value of a typical year of life and the value of a year of life in a state of severe dependence (0.55 QALYs). It is also assumed that this individual suffers a decreased length of life with a life expectancy of 50 years. It is assumed that members of the AIDS subpopulation do not have offspring; hence, this cost is assumed not to apply in their case.

- Mild GI illness: For the general population, the cost of a year of GI illness is assumed to be the difference between the value of a typical year of life the value of a year of life in severe pain (0.64 QALYs). A mild illness is assumed to last one day and hence has a cost equal the cost of a year of illness divided by 365 days. For members of the AIDS subpopulation, this cost is assumed to equal one-half the average cost for members of the general population.

- Moderate to severe illness: This condition is assumed to last 14 days; its cost is therefore assumed to be 14 times the cost of a mild GI illness.

- Death due to microbial infection: The QALY cost of this event is the value of the life years lost.

Table 5-25 summarizes the QALY costs for these health endpoints for the general population. Specifically, the QALY costs in Table 5-25 represent the net present value of the lost QALYs due the health effect at the time the effect occurs. Table 5-26 lists the corresponding values for the AIDS subpopulation. Note that although the QALY cost is the same for death resulting from cancer and death resulting from microbial infection, the latter is, in practice, of

Table 5-25
Age-Specific QALY Costs for All Health Endpoints: General Population
Discount Rate of 3%

Age at Which Event Occurs	Cancer Illness	Cancer Death	1 Year Infertility	Develop. Defect	Mild GI Illness (1 Day)	Moderate to Severe GI Illness (14 days)	Microbe-Induced Death
0 to 4	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	16.211
5 to 9	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	16.224
10 to 14	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	16.108
15 to 19	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	15.951
20 to 24	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	15.775
25 to 29	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	15.546
30 to 34	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	15.233
35 to 39	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	14.821
40 to 44	0.900	16.211	0.120	14.839	8.49E-4	8.49E-4	14.275
45 to 49	0.842	16.211	0.120	14.839	7.70E-4	7.70E-4	13.546
50 to 54	0.842	16.211	0.120	14.839	7.70E-4	7.70E-4	12.777
55 to 59	0.746	16.211	0.120	14.839	6.38E-4	6.38E-4	11.792
60 to 64	0.746	16.211	0.120	14.839	6.38E-4	6.38E-4	10.882
65 to 69	0.674	16.211	0.120	14.839	5.40E-4	5.40E-4	9.798
70 to 74	0.674	16.211	0.120	14.839	5.40E-4	5.40E-4	8.730
75 to 79	0.630	16.211	0.120	14.839	4.79E-4	4.79E-4	7.464
80 to 84	0.630	16.211	0.120	14.839	4.79E-4	4.79E-4	6.037
85+	0.616	16.211	0.120	14.839	4.60E-4	4.60E-4	4.040

Table 5-26
Age-Specific QALY Costs for All Health Endpoints: AIDS Subpopulation
Discount Rate of 3%

Age at Which Event Occurs	Cancer Illness	Cancer Death	1 Year Infertility	Develop. Defect	Mild GI Illness (1 Day)	Moderate to Severe GI Illness (14 days)	Microbe-Induced Death
0 to 4	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
5 to 9	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
10 to 14	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
15 to 19	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
20 to 24	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
25 to 29	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
30 to 34	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
35 to 39	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
40 to 44	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
45 to 49	0.200	4.820	NA[a]	NA[a]	3.60E-4	5.04E-3	4.820
50 to 54	0.200	4.818	NA[a]	NA[a]	3.60E-4	5.04E-3	4.818
55 to 59	0.200	4.815	NA[a]	NA[a]	3.60E-4	5.04E-3	4.815
60 to 64	0.200	4.806	NA[a]	NA[a]	3.60E-4	5.04E-3	4.806
65 to 69	0.200	4.783	NA[a]	NA[a]	3.60E-4	5.04E-3	4.783
70 to 74	0.200	4.722	NA[a]	NA[a]	3.60E-4	5.04E-3	4.722
75 to 79	0.200	4.561	NA[a]	NA[a]	3.60E-4	5.04E-3	4.561
80 to 84	0.200	4.133	NA[a]	NA[a]	3.60E-4	5.04E-3	4.133
85+	0.200	3.000	NA[a]	NA[a]	3.60E-4	5.04E-3	3.000

Notes: *Costs for these endpoints are assumed not to be applicable to the AIDS subpopulation since it is assumed that members of this group do not bear children.*

greater consequence because there is effectively no latency period in the case microbial illness. Cancer, on the other hand, is generally manifest many years after an exposure; its cost must therefore be appropriately discounted to reflect this delay.

The uncertainty associated with these estimates has not been quantified, even though it is clear that these values are not known precisely. To assess the potential importance of this uncertainty, it is assumed that the plausible range of QALY cost values for each health endpoint is log-uniformly distributed between half the point estimate derived in this section, and twice this point estimate[18].

[18] A quantity is log-uniformly distributed between a and b if the log of its value is uniformly distributed between log(a) and log(b).

Table 5-10
Cancer Slope Factor Distributions

Chemical	Cancer Class	MLE Slope per mg/kg-d	Upper 95% Slope per mg/kg-d	Lognormal Mean	Lognormal Standard Deviation	
CHBrCl2	B2	5.7E-03	6.2E-02	-5.2	1.5	Renal adenomas and adenocarcinomas
CHBr2Cl	C	7.2E-04	8.4E-02	-7.2	2.9	Hepatocellular adenomas and adenocarcinomas
CHBr3	B2	3.4E-04	7.9E-03	-8.0	1.9	Neoplastic lesions in large intestine
CH	C	4.1E-02	1.3E-01	-3.2	0.7	Hepatocellular adenomas and adenocarcinomas
DCA	B2	1.4E-03	1.0E-01	-6.6	2.6	Hepatocellular adenomas and adenocarcinomas
TCA	C	4.9E-02	8.4E-02	-3.0	0.3	Liver neoplasms
Bromate	B2	3.2E-01	4.9E-01	-1.1	0.3	Renal adenomas and adenocarcinomas

Note: Chloroform is not in the risk estimate; it is considered a threshold carcinogen. TOX levels are below threshold.

CHCl3	B2	3.1E-03	6.1E-03	-5.8	-5.1	Renal tumors

Table 5-11

Developmental Slope Factor Distributions

Chemical	MLE Slope, Exposure > Threshold per mg/kg-d	95% Upper Slope, Exp > Threshold per mg/kg-d	Lognormal Mean	Lognormal Standard Deviation	
DCA	8.6E-03	1.3E-02	-4.8	0.3	Visceral malformations Total
TCA	2.0E-02	3.0E-02	-3.9	0.2	Fetal body weight - male
MBA	8.4E-03	2.3E-02	-4.8	0.6	Fetal crown rump length
DBA	8.6E-03	1.3E-02	-4.8	0.3	Estimated using DCA as surrogate
BCA	8.6E-03	1.3E-02	-4.8	0.3	Estimated using DCA as surrogate
DCAN	5.4E-02	1.6E-01	-2.9	0.7	Visceral malformations cardiovascular
TCAN	2.1E-01	3.4E-01	-1.6	0.3	Visceral malformations Total
BCAN	1.6E-01	2.4E-01	-1.8	0.2	Visceral malformations Total
DBAN	2.1E-01	3.4E-01	-1.6	0.3	Estimated using TCAN as surrogate

Note: CHBrCl2 and MCA are not in the risk estimate; the model estimated a threshold above exposure levels.

CHBrCl2	4.0E-02	3.1E-02	-3.2	-0.2	
MCA	9.0E-05	6.0E-03	-9.3	2.6	

Table 5-12

Reproductive Slope Factor Distributions

Chemical	MLE Slope Exposure > Threshold per mg/kg-d	95% Upper Slope, Exp > Threshold per mg/kg-d	Lognormal Mean	Lognormal Standard Deviation	Effect
DBA	2.5E-02	6.0E-02	-3.7	0.5	Number cauda sperm
DCA	2.5E-02	6.0E-02	-3.7	0.5	Estimated using DBA as surrogate
BCA	2.5E-02	6.0E-02	-3.7	0.5	Estimated using DBA as surrogate

Table 5-22
Conditional Survival Probabilities for the General Population[a,b]

	0 to 4	5 to 9	10 to 14	15 to 19	20 to 24	25 to 29	30 to 34	35 to 39	40 to 44	45 to 49	50 to 54	55 to 59	60 to 64	65 to 69	70 to 74	75 to 79	80 to 84	85
0 to 4	1.000																	
5 to 9	0.991	1.000																
10 to 14	0.990	0.999	1.000															
15 to 19	0.989	0.998	0.999	1.000														
20 to 24	0.985	0.994	0.995	0.996	1.000													
25 to 29	0.979	0.988	0.989	0.991	0.995	1.000												
30 to 34	0.973	0.982	0.983	0.985	0.989	0.994	1.000											
35 to 39	0.966	0.975	0.976	0.977	0.981	0.986	0.992	1.000										
40 to 44	0.956	0.964	0.965	0.967	0.971	0.976	0.982	0.990	1.000									
45 to 49	0.942	0.951	0.952	0.953	0.957	0.962	0.968	0.976	0.986	1.000								
50 to 54	0.925	0.933	0.934	0.935	0.939	0.944	0.950	0.958	0.968	0.981	1.000							
55 to 59	0.899	0.907	0.908	0.909	0.913	0.918	0.923	0.931	0.941	0.954	0.972	1.000						
60 to 64	0.861	0.868	0.869	0.870	0.874	0.879	0.884	0.891	0.901	0.913	0.930	0.957	1.000					
65 to 69	0.803	0.810	0.811	0.812	0.815	0.820	0.825	0.831	0.840	0.852	0.868	0.893	0.933	1.000				
70 to 74	0.724	0.731	0.731	0.732	0.735	0.739	0.744	0.750	0.758	0.768	0.783	0.805	0.841	0.902	1.000			
75 to 79	0.618	0.624	0.625	0.625	0.628	0.631	0.635	0.640	0.647	0.656	0.668	0.688	0.718	0.770	0.854	1.000		
80 to 84	0.487	0.491	0.492	0.492	0.494	0.497	0.500	0.504	0.509	0.517	0.526	0.542	0.566	0.606	0.672	0.787	1.000	
85+	0.332	0.335	0.335	0.335	0.337	0.339	0.341	0.343	0.347	0.352	0.359	0.369	0.385	0.413	0.458	0.536	0.681	

Notes: [a]*Computed from values listed in the second column of Table 5-21*

[b]*Entries represent the probability of surviving until the beginning of the age range listed in the far left column of an entry's row assuming that an individual is alive at the beginning of the age range listed at the head of that entry's column. For example, an individual who has survived until age 5 (the beginning of the 5 to 9 age range) has an 86.8% probability of surviving until age 60 (the beginning of the age range from 60 to 64). These probabilities are interpreted as being the probability of surviving until the age range listed in the left column of an entry's row given survival to the age range listed at the head of the column listing that entry.*

Table 5-23
Conditional Survival Probabilities for the AIDS Subpopulation[a,b]

	0 to 4	5 to 9	10 to 14	15 to 19	20 to 24	25 to 29	30 to 34	35 to 39	40 to 44	45 to 49	50 to 54	55 to 59	60 to 64	65 to 69	70 to 74	75 to 79	80 to 84	85+
0 to 4	1.0E+0																	
5 to 9	5.3E-1	1.0E+0																
10 to 14	2.8E-1	5.3E-1	1.0E+0															
15 to 19	1.5E-1	2.8E-1	5.3E-1	1.0E+0														
20 to 24	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0													
25 to 29	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0												
30 to 34	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0											
35 to 39	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0										
40 to 44	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0									
45 to 49	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0								
50 to 54	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0							
55 to 59	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0						
60 to 64	5.0E-4	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0					
65 to 69	2.7E-4	5.0E-4	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0				
70 to 74	1.4E-4	2.7E-4	5.0E-4	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0			
75 to 79	7.5E-5	1.4E-4	2.7E-4	5.0E-4	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0		
80 to 84	4.0E-5	7.5E-5	1.4E-4	2.7E-4	5.0E-4	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	1.0E+0	
85+	2.1E-5	4.0E-5	7.5E-5	1.4E-4	2.7E-4	5.0E-4	9.4E-4	1.8E-3	3.3E-3	6.3E-3	1.2E-2	2.2E-2	4.2E-2	7.9E-2	1.5E-1	2.8E-1	5.3E-1	

Notes: [a]*Computed from values listed in the third column of Table 5-21*

[b]*Entries represent the probability of surviving until the beginning of the age range listed in the far left column of an entry's row assuming that an individual is alive at the beginning of the age range listed at the head of that entry's column. For example, an individual who has survived until age 25 (the beginning of the 25 to 29 age range) has a 53% probability of surviving until age 30 (the beginning of the age range from 30 to 34). These probabilities are interpreted as being the probability of surviving until the age range listed in the left column of an entry's row given survival to the age range listed at the head of the column listing that entry.*