United States
Environmental Protection Agency

Office of Water (WH-550)
Washington DC 20460

EPA 814-R-92-007
August 1992

EPA

Simulation of Microbial Occurrence, Exposure, and Health Risks After Drinking Water Treatment Processes

THIS PAGE INTENTIONALLY BLANK

SIMULATION OF MICROBIAL OCCURRENCE, EXPOSURE AND HEALTH RISKS AFTER DRINKING WATER TREATMENT PROCESSES

WILLIAM D. GRUBBS, BRUCE A. MACLER* and STIG REGLI*

Science Applications International Corporation

and

*U.S. Environmental Protection Agency

ABSTRACT

For the development of the Disinfectant/Disinfection Byproduct Rule, EPA wishes to compare human health risks from microbial infection with those from chemical disinfectants and their byproducts. A direct comparison using available data is not possible at this time. Therefore, EPA is approaching this problem with the use of computer models that simulate occurrence levels of pathogenic organisms in raw water, then simulate disinfection and production of disinfection byproducts. The microbial and chemical concentrations thus generated are then used to estimate potential health risks. This paper presents the methodology used for these simulations and estimations and discusses the assumptions and uncertainties inherent to this modeling process.

Two distinct sources of variation were examined in this analysis. Summary measurements of existing data for Giardia occurrence from different cities reflected a geographic variation. Measurements from the same city but on different days reflected a temporal variation. These variations were characterized from data collected by Hibler (1988) and LeChevallier, et al (1991). The lognormal distribution was used to describe the geographic variation, and a combination of two discrete distributions, the delta and the negative binomial distributions, was used to describe the temporal variation.

Annual averages of Giardia occurrence in raw surface water for 100 cities were simulated based on the geographic variation. These averages in raw surface water were the basis for input to a simulation model, where treatment was applied as a function of the raw surface water quality. This simulation model used engineering and chemical equations to predict Giardia in finished surface water.

Giardia occurrence in finished surface water also exhibits a geographic and temporal variation. These types of variation were employed in conjunction with a dose-response function that related the probability of infection to the number of Giardia cysts in finished surface water. Quantities related to this function were used to estimate endemic levels and the frequency of an outbreak. Additional refinements to this analysis were performed to examine the effects of secondary infection and system malfunctions on the results based on Giardia occurrence in finished surface water.

INTRODUCTION

EPA is developing National Primary Drinking Water Standards for various chemical disinfectants and their disinfection byproducts. The goals of this Disinfectant/Disinfection Byproducts (D/DBP) Rule are to ensure that drinking water is microbiologically safe at any limits set for disinfectants and byproducts, and that the disinfectants and byproducts themselves do not pose unacceptable risks at these limits. EPA's approach in developing this rule is to consider different regulatory scenarios that achieve different definitions of microbial safety and risk levels from disinfectants and byproducts. These risks are linked, in that any increase in disinfection to lower microbial risks requires that use of more disinfectants and consequently yields higher levels of byproducts, thus increasing chemical health risks. Determination of the magnitude of microbial and disinfectant/byproduct risks as a function of different water treatment trains and source water qualities is essential to crafting a rule that will minimize overall health risks from drinking water.

The comparison of microbial health risks with those generated from drinking water treatment for given treatments is not directly possible using currently available data. As a result, EPA is approaching this problem with the use of computer models that simulate the occurrence levels of pathogenic organisms in raw water, then simulate disinfection and production of certain disinfection byproducts of health concern. The microbial and chemical concentrations generated for this "treated" water are then used to estimate potential health risks.

This paper presents the methodology used for these simulations and estimations and discusses the assumptions and uncertainties inherent to this modeling process. Giardia lamblia was selected as the target organism for the modeling effort since a) the existing data base for its occurrence is the most extensive of any pathogenic microorganism found in drinking water; b) CT values have been developed for predicting disinfection inactivation efficiencies; c) it is much more resistent to disinfection than most other waterborne pathogens and therefore changes in disinfection practice are more likely to affect Giardia exposures than those for most other pathogens; and d) dose-response data are available for Giardia for estimating risk from exposure.

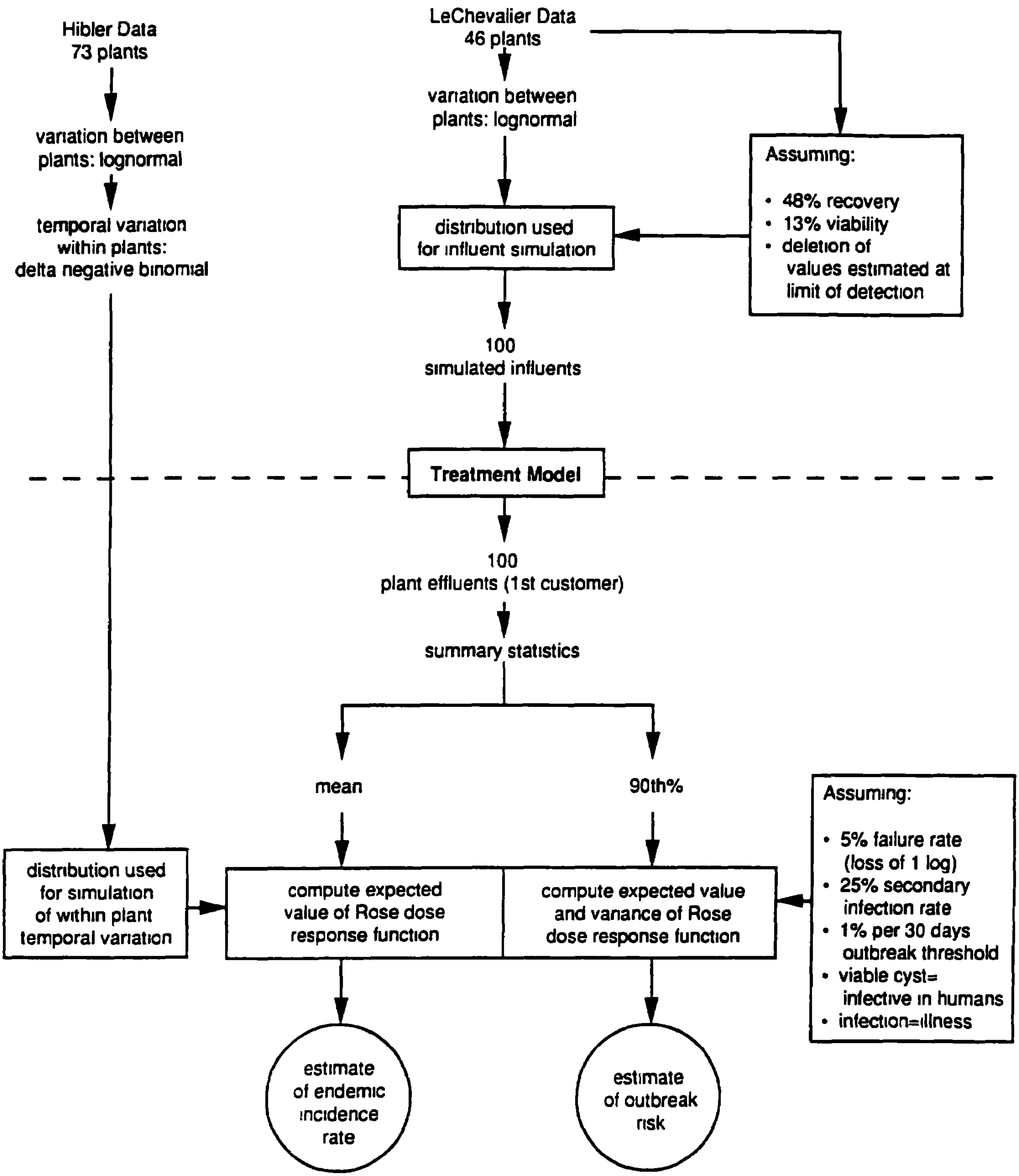
METHODOLOGY AND DATA

A flowchart describing the process of estimating endemic levels and outbreak frequency is presented in Figure 1. Details of the methodology and assumptions follow.

Geographic variation of Giardia occurrence in raw surface water data

Data collected by LeChevallier, et al (1991) were used to characterize the geographic variation of Giardia cyst concentrations in source waters for different cities. These data were used to represent the annual average cyst concentrations for different cities. These data are not appropriate for assessing the temporal variation described below (i.e., the changes in the number of Giardia cysts over some time period). The listing of the 85 measurements in this data base is given in Appendix A. The 15 measurements with a '*' under the "OBSERVATION WAS DELETED" column were not included in the analysis because these observations were estimated Giardia levels based on the detection limit, rather than actual observations. Each measurement was

Overview of *Giardia* Modelling



multiplied by 2.08 to reflect a retrieval efficiency (i.e., the measurement technique was not able to observe the true number of cysts present) and by 0.13 as a viability factor; that is, cysts that are counted, but because of morphological characteristics are not considered actually alive. Cities with multiple measurements were averaged to obtain one measurement for each city. These 46 city averages are also presented in Appendix A, along with the number of annual averages from which these city averages were constructed.

Data collected by Hibler (1988) also provide information on the geographic variation, if, for example, the arithmetic average of the daily measurements within a city is taken to construct a city average. However, it was determined that the LeChevallier data were more appropriate for characterizing the geographic variation because the methods used in the collection of the LeChevallier data were more advanced, and this data encompassed a broader geographic region.

Distributional assumptions

The 46 city averages based on the LeChevallier, et al (1991) data were tested for normality and lognormality, assuming independence among cities. The Shapiro-Wilk (Shapiro and Wilk, 1965) test, as calculated in the SAS® procedure PROC UNIVARIATE (SAS, 1990) was chosen to test the hypotheses of normality and lognormality. Lognormality was examined by testing the natural logarithm of the 46 city averages for normality. The results of this analysis are shown in Appendix A. The null hypothesis of normality of the city averages was soundly rejected ($p<0.001$), but the null hypothesis of lognormality could not be rejected ($p=0.582$). Consequently the lognormal distribution was used to characterize the geographic variation, and a lognormal distribution was used to generate the input into the simulation model, which predicted Giardia in finished surface water and is discussed in further detail below.

Although the Hibler data was not used in characterizing geographic variation, an inspection of the 73 city averages constructed from this data also support the hypothesis of lognormality of the city averages.

Temporal variation of Giardia occurrence in raw surface water data

Data collected by Hibler (1988) were used to characterize the temporal variation of Giardia cyst occurrence in raw surface water. A summary of 1,515 measurements across 73 cities is given in Appendix B. The number of cysts was also multiplied by 2.0 to reflect a retrieval efficiency, as discussed by Hibler (1988), and rounded to the nearest integer, since the nature of the measurement (number of Giardia cysts) is inherently discrete. No information was available on the exact date of sampling to assess any correlation among measurements across time. The distributional methodology was consequently developed based on an assumption of independence among daily measurements. Cities were also assumed to be independent. Selected summary statistics on the Hibler data are provided in Appendix B. An important statistic from this data, which was crucial in the distributional assumption for characterizing the temporal variation of Giardia cyst occurrence, involves the large number of zero Giardia cyst measurements. Approximately 72 percent (1,087/1,515) of all Giardia cyst measurements in this data base are 0.

Distributional assumptions

Typical discrete distributions for modeling the number of Giardia cysts include the Poisson distribution and the negative binomial distribution. If the random variable X represents the number of Giardia cysts in raw surface water, then under the Poisson distribution,

$$Pr(X=x) = \frac{\exp(-\lambda)\cdot\lambda^x}{x!}, \qquad \lambda>0, \qquad x=0,1,2,\ldots$$

Under the negative binomial distribution,

$$Pr(X=x) = \frac{\Gamma(q+x)}{\Gamma(q)\cdot x!}\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^x, \qquad q>0,\ m>0,\ x=0,1,2,\ldots$$

where $\Gamma(\cdot)$ is the gamma function described in mathematical tables handbooks, such as the CRC handbook (Beyer, 1968) and $x!=x\cdot(x-1)\cdot(x-2)\cdot\ldots1$. The λ parameter for the Poisson distribution and the parameters m and q for the negative binomial distribution are usually estimated from available data.

A characteristic of the Poisson distribution is that the mean and variance are equal. The variance is larger than the mean for the negative binomial distribution. In particular, if E(X) represents the mean of a distribution of a random variable X, and V(X) represents the variance of X, then for the Poisson distribution,

$$E(X) = \lambda \text{ and } V(X) = \lambda.$$

For the negative binomial distribution,

$$E(X) = m \text{ and } V(X) = m+(m^2/q).$$

In analyzing the data, temporal distributions of Giardia cyst occurrence in raw surface water were developed separately on a city-by-city basis. It was clear that neither the Poisson nor the negative binomial distribution in their present form was able to adequately represent the data, because of the large number of zeroes. A value of zero is permissible in both the Poisson and negative binomial distributions, but to choose parameters that will adequately model the high percentage of zeroes in this data would cause the probability of larger values occurring to be extremely small. Also an inspection of the data reveals a larger variance than mean for many of the cities. This aspect of the data favors the choice of the more flexible negative binomial distribution over the Poisson distribution, which has an equal mean and variance.

Consequently, a modification of the negative binomial distribution was used, whereby a point distribution placed at zero was combined with the negative binomial distribution. The point distribution is often referred to as the delta distribution, and the mixture of these two distributions will subsequently be referred to as the delta-negative binomial (DNB) distribution. The form of the DNB distribution, where the random variable X represents the number of Giardia cysts in raw surface water, is

$$Pr(X=x) = \delta \cdot I_0(x) + (1-\delta) \cdot \frac{\Gamma(q+x)}{\Gamma(q) \cdot x!} \cdot \left(\frac{q}{q+m}\right)^q \cdot \left(\frac{m}{q+m}\right)^x, \qquad 0<\delta<1,\ q>0,\ m>0,\ x=0,1,2,\ldots$$

where $I_0(x)=1$ if $x=0$ and 0 otherwise.

The mean, E(X), and the variance, V(X), of the DNB distribution are

$$E(X) = (1-\delta) \cdot m, \text{ and}$$
$$V(X) = (1-\delta) \cdot m \cdot [(\delta \cdot m)+1+(m/q)]$$

The derivations of these quantities are given in Appendix C.

Three parameters (δ, m, and q) need to be estimated for a DNB distribution. Parameter estimates for 42 cities in the Hibler data base are presented in Appendix D. All cities in this appendix had at least four measurements; the other cities not included in this appendix had less than four measurements. The cities are listed in descending order by sample size. These parameter estimates for each city were derived through the SAS® procedure PROC NLIN. The starting values (δ_s, m_s, and q_s) used were

$\delta_s = n_0/n$
m_s = meanpos
q_s = meanpos^2/(varpos-meanpos) if varpos>meanpos,
= ∞ if varpos$\leq$meanpos,

where n_0 = number of zero cyst measurements,
n = number of measurements,
meanpos = mean of the nonzero cyst measurements,
varpos = variance of the nonzero cyst measurements, and
$q_s = \infty$ was replaced by a large number (100,000) for computational purposes.

A graphical illustration of the actual data (72 observations) and a DNB distribution based on the estimated parameters for city 109 is given in Appendix E.

Relating *Giardia* occurrence in raw surface water to finished water

A lognormal distribution based on the LeChevallier, et al (1991) data was the basis for input into a simulation model that predicted *Giardia* in finished surface water. Based on the 46 city averages described previously, a logmean of 5.39 and a logstandard deviation of 1.67 were used as input parameters (see Appendix A for the derivation of these quantities) to the model, and 100 city averages were generated with the SAS® function RANNOR. The simulation model utilized engineering and chemical equations to apply treatment as a function of raw water quality and generated 100 city averages for *Giardia* in finished water as output. (See Gelderloos, et al (1992) and Cromwell et al (1992) for a complete description of this simulation model.) In general, this model simulated conventional water treatment of surface water without lime softening (i.e., coagulation, sedimentation, filtration and chemical disinfection), capable of meeting Surface Water Treatment Rule (SWTR) requirements, with and without the use of an alternative disinfectant to chlorine. In addition, an "enhanced" SWTR level of treatment was simulated, using EPA SWTR guidance for poorer quality waters, to consider the potential value of increasing the level of disinfection as a function of poorer source water quality. Treatment performance under the enhanced SWTR was estimated by assuming that filtration achieved 2.5 log removal of *Giardia* cysts and that disinfection achieved additional log inactivation of cysts predicted by CT equations. The enhanced SWTR specified a level of treatment necessary to achieve an approximately constant average *Giardia* cyst concentration at the first customer (Gelderloos, et al, (1992)).

Assuming that the removal efficiency for *Giardia* cysts between raw and finished surface water is binomial and the distribution of *Giardia* cysts in raw water is a DNB distribution, then the distribution of *Giardia* cysts in finished surface water is also a DNB distribution. The assumption of a binomial removal process assumes a constant probability p of survival of a *Giardia* cyst between raw and finished surface water. The mathematical proof behind this statement is included in Appendix F. The log removal is directly related to this probability p; in fact, log removal = $-\log_{10}(p)$. As seen by this proof, if the number of *Giardia* cysts in raw water exhibits a DNB distribution with parameters δ, m, and q, then the number of *Giardia* cysts in finished water has a DNB distribution with parameters δ, $m \cdot p$, and q. In particular, if the random variable Y represents the number of *Giardia* cysts in finished surface water, then the mean, E(Y), and the variance, V(Y), are

$$E(Y) = (1-\delta)\cdot m\cdot p \qquad (1)$$
$$\text{and } V(Y) = (1-\delta)\cdot m\cdot p\cdot[(\delta\cdot m\cdot p)+1+((m\cdot p)/q)]$$

A DNB distribution was consequently assumed for each of the 100 cities, using values for δ and q as derived from the Hibler (1988) data, along with the annual averages as output from the simulation model. A value for δ of 0.7 was chosen, very similar to the approximately 72% of zero measurements found in the Hibler data. A value for q of 4.5 was selected to represent a typical estimate for q from the cities in Hibler's data. Since $E(Y) = (1-\delta)\cdot m\cdot p$ from formula (1), $m \cdot p$ was estimated as the annual average divided by $(1-\delta)$. A DNB distribution with these values of δ, $m \cdot p$, and q were consequently used, along with the dose-response relationship described below, to estimate endemic levels and outbreak frequency.

In subsequent results, certain summary statistics from the 100 cities were used to characterize the distribution. In particular, for estimating endemic levels, results were based on the arithmetic average of the number of cysts in finished water across the 100 cities. For estimating outbreak frequency under nominal SWTR conditions, results were based on the number of cysts at the city closest to the 90th percentile of the 100 cities. Under an enhanced SWTR, no outbreaks should occur by definition, since the level of treatment should ensure that the infection rate would be well below the assumed outbreak threshold of greater than one percent of the population becoming infected within a one-month period. A small number of these 100 cities were not included in the analysis because the requirements of the SWTR and taste and odor constraints could not be met (see Gelderloos, et al (1992) for a further description of the deletion of cities from the Giardia analysis).

Dose-response relationship

Using a risk assessment model from Rose, et al (1991) based on human infectivity studies (Rendtorff, 1954; Rendtorff and Holt, 1954), a dose-response relationship was developed to estimate the risk of infection due to waterborne exposure to Giardia. In the Rendtorff studies, a total of 40 volunteers were fed Giardia cysts in capsules, and a positive response was measured by cyst excretion in the feces. Infection was the measure of a positive response and not illness. While no infection resulted in illness in the Rendtorff study, which used healthy male prisoners, we used the conservative assumption for the outbreak analysis that all infections result in illness. This appears reasonable, since a substantial number of infected individuals do become ill, as indicated by over 100 reported waterborne outbreaks of giardiasis in the U.S. since 1965. Also, the paper by Regli, et al (1991), which compared predicted infection rates with actual illness rates in communities with waterborne outbreaks of giardiasis, appears to support this assumption as being correct within one order of magnitude.

An exponential dose-response function was used to relate the probability of infection to the number of Giardia cysts ingested. If the random variable D represents the number of Giardia cysts ingested and the random variable P_D represents the probability of an infection, then the exponential dose-response function can be related to the probability of infection as

$$P_D=1-\exp(-r \cdot D),$$

where r is a parameter estimated from the data. In this case, using data from the Rose, et al (1991), a value for r of 0.02 was derived. The data from Rose, et al, are presented in Appendix G, along with the estimation of the parameter r according to the exponential dose-response function using SAS® PROC NLIN.

Estimating endemic levels

The dose-response function and the DNB distribution were combined to estimate the endemic level of infection in the population. The endemic level measures the frequency of occurrence of disease in a population that exists on an ongoing basis. In particular for this analysis, the estimate of endemic levels was expressed as the expected number of infections per 10,000 per year. A "case" was defined as an infection episode; that is, one person can have more than one infection per year.

A key assumption of this analysis regards the length of infection. It is assumed that once a person is infected, that person remains infected for a 30-day period. That is, if a person were infected on day 1 then he would remain infected until day 30, and not be infected on day 31. This assumption is aimed at eliminating implausible co-occurring infections resulting from the probabilistic approach taken here.

The algorithm for estimating the average number of cases per 10,000 per year is described below:

1) The units of the Hibler (1988) and LeChevallier, et al (1991) data are in cysts per 100 gallons, and the parameters developed from the Hibler data for use in this analysis are based on the number of Giardia cysts per 100 gallons. Consequently an outcome chosen from a finished surface water DNB distribution will be expressed in cysts per 100 gallons. Convert the dosage from 100 gallons to 2 liters and calculate the probability of an infection; that is,

$$\begin{aligned} P_D^* &= 1-\exp(-r\cdot D\cdot 0.005283) \\ &= 1-\exp(-r^*\cdot D), \end{aligned} \qquad (2)$$

where 2 liters $=0.005283\cdot 100$ gallons and $r^* = r\cdot 0.005283$.

2) Determine the average probability of an infection, which is the expected value of the dose-response function in (2), assuming D is an outcome from the DNB distribution with parameters δ, $m\cdot p$, and q. This expected value is

$$=(1-\delta)\left\{1-\left[\frac{q}{q+(m\cdot p)\cdot(1-e^{-r^*})}\right]^q\right\} \qquad (3)$$

The derivation of this quantity is found in Appendix H.

3) Determine the probability of a person being infected for z days out of a year, where z can range between 0 and 365. The probability of being infected on any given day, given that a person is not infected, is assumed to be equal to the average probability of infection in equation (3). The general formula for a person being infected z days out of a year is

$$Pr(Z=z) = \frac{(365-z+t)!}{(365-z)! \bullet (t)!} \bullet exdr^{v} \bullet (1-exdr)^{365-z} \qquad z=0,1,\ldots 365$$

where t = largest integer less than or equal to (i/30),
exdr = average probability of infection given in (3), and
v = smallest integer greater than or equal to (i/30).

The values t and v were accomplished using the SAS® functions FLOOR and CEIL, respectively. Specific examples of how this formula was derived are in Appendix I.

4) Translate the number of days being infected in a year to the number of cases. If a person is infected for 0 days in a year, that person has 0 cases for the year. If a person is infected between 1 and 30 days in a year, then that person has 1 case for the year. Two cases in a year are for 31 to 60 days infected, and so on, up to 13 cases for between 360 and 365 days in a year.

5) Calculate the expected number of cases, which is the sum from 0 to 365 of the number of cases corresponding to day z times the probability of being infected z days. Multiply the result by 10,000.

As additional modifications to this analysis, two extra issues were investigated. In particular, the influence of a secondary infection rate and a reduction in system effectiveness were studied.

A secondary infection rate assumes that a person being infected can pass giardiasis to another person a certain percentage of the time. Various alternatives can be studied; in the results presented subsequently, a secondary infection rate of 25% was assumed. Using the principles of the geometric series

$$1 + a + a^2 + a^3 + \ldots = 1/(1-a)$$

a 25% secondary infection rate translates to an increase in the endemic rate by 33%. The results for estimating endemic rates as described above were adjusted accordingly to reflect a 25% secondary infection rate.

The consequences of a reduction in system effectiveness were also studied. In particular, subsequent results assumed that 5% of the time the removal for a given city was 1 log (factor of 10) less than the nominal removal. The nominal removal was calculated from the annual average for the raw and finished water at a given city. For the purposes of this analysis, time is treated as a continuum; that is, the system had a 1 log reduction in effectiveness 5% of the time scattered over a period (the reduction is not clustered on any particular day, for example). The nominal removal is calculated from the annual average for raw and finished water at a given city.

In particular, let

RAW = average number of cysts in raw surface water at a particular city,
FIN_{05} = average number of cysts in finished surface water at a particular city assuming a reduction in system effectiveness 5% of the time,
FIN_R = average number of cysts in finished surface water at a particular city assuming no reduction in system effectiveness for a given (R) log removal from treatment, and
FIN_{R-1} = average number of cysts in finished surface water at a particular city assuming a reduction in system effectiveness of 1 log.

Then

$$FIN_{05} = 0.95 \bullet FIN_R + 0.05 \bullet FIN_{R-1}$$

$$= 0.95 \bullet \left(\frac{RAW}{10^R}\right) + 0.05 \bullet \left(\frac{RAW}{10^{R-1}}\right)$$

$$= 0.95 \bullet \left(\frac{RAW}{10^R}\right) + 0.05 \bullet 10 \bullet \left(\frac{RAW}{10^R}\right)$$

$$= 1.45 \bullet \left(\frac{RAW}{10^R}\right)$$

$$= 1.45 \bullet FIN_R$$

Consequently, a 5% reduction in system effectiveness corresponds to a 45% increase in the average number of cysts in finished surface water at a given city. The results for estimating endemic rates as described above were adjusted accordingly to reflect this 5% reduction in system effectiveness.

Estimating outbreak frequency

The dose-response function and the DNB distribution were also combined to estimate the frequency of an outbreak in the population. For the purposes of this analysis, an outbreak has been defined as observing giardiasis infection in greater than one percent of the population within any given 30-day period. This assumption is based on observations that the awareness of waterborne disease often does not occur unless at least one percent of the population becomes ill within a time frame of about a month (Regli, et al, 1991). The population considered in the analysis here was defined as first customers closest to the treatment plant with consequently minimal distribution system CT disinfection. The assumption that outbreaks are identified by a one percent infection rate at the first customer versus one percent for the entire population is somewhat arbitrary and may be overly conservative, depending upon the relative population density near the first customer.

For this analysis, the estimate of the frequency of an outbreak is related to the probability that an outbreak occurs. The results were based on the simulated number of cysts closest to the upper 90th percentile of the modeled distribution.

The algorithm for estimating the outbreak frequency is described below:

1) Determine the average probability of an infection and the variance in the probability of an infection. The average probability of an infection is the expected value of the dose-response function in equation (2), and as given in equation (3), is

$$exdr = (1-\delta)\left\{1-\left[\frac{q}{q+(m \bullet p)\bullet(1-e^{-r^{*}})}\right]^{q}\right\}$$

The variance in the probability of an infection is the variance of the dose-response function in equation (2), and is

$$\begin{aligned} vardr = (1-\delta)\Bigg\{ \delta \left[1 - 2\left(\frac{q}{q+((m \bullet p)\bullet(1-e^{-r^{*}}))}\right)^{q}\right] + \\ \left(\frac{q}{q+((m \bullet p)\bullet(1-e^{-(2\bullet r^{*})}))}\right)^{q} - (1-\delta)\left(\frac{q}{q+((m \bullet p)\bullet(1-e^{-r^{*}}))}\right)^{2q}\Bigg\} \end{aligned} \quad (4)$$

The derivations of these quantities are found in Appendix H.

2) Let I denote the event of being infected on a given day, and I_{30} the event of being infected in any 30-day period. Being infected in any 30-day period can be considered as the sum of 30 consecutive events of being infected on a given day. Expressed mathematically, the probability that an outbreak occurs is expressed as

$$\Pr(I_{30} > 0.01),$$

Calculate the average probability of an infection in a 30-day period = $exdr_{30}$ = $30 \cdot exdr$. Calculate the variance in the probability of an infection in a 30-day period = $vardr_{30}$ = $30 \cdot vardr$. The formulas for exdr and vardr are given in equations (3) and (4), respectively.

3) The dose-response function in equation (2) is an exponential function of X and X has a DNB distribution. Consequently, the dose-response function will not have a normal distribution. But by the Central Limit Theorem, it is assumed that the mean (or sum, in this analysis) of 30 samples (I_{30}) of the dose-response function is normally distributed with mean $exdr_{30}$ and variance $vardr_{30}$. Consequently, using the Central Limit Theorem,

$$Pr(I_{30} > 0.01) =$$

$$Pr\left(\frac{I_{30} - exdr_{30}}{\sqrt{vardr_{30}}} > \frac{0.01 - exdr_{30}}{\sqrt{vardr_{30}}}\right) =$$

$$1 - \Phi\left(\frac{0.01 - exdr_{30}}{\sqrt{vardr_{30}}}\right) \qquad (5)$$

where $\Phi(\cdot)$ represents the cumulative distribution function of the standard normal distribution. This was performed using the SAS® function PROBNORM.

4) The quantity in equation (5) is an estimate of the probability of an outbreak in any 30-day period. To estimate the expected number of outbreaks in a year, multiply this probability by 365. The number 365 is used to estimate a years' worth of 30-periods (for example, days 1-30, days 2-31, days 3-32, and so on). To estimate the expected number of years to an outbreak, take the reciprocal of the expected number of outbreaks in a year.

The impacts of a 25% secondary infection rate and a 5% reduction in system effectiveness were also incorporated into the estimates of outbreak frequency. The threshold for an outbreak was changed from 1% to 0.75% to account for a 25% secondary infection rate and the average number of cysts in finished water was increased by 45% to reflect a 5% reduction in system effectiveness. The subsequent results for estimating outbreak frequency reflect the 25% secondary infection rate and 5% reduction in system effectiveness.

RESULTS AND DISCUSSION

Table 1 presents estimates of endemic levels and outbreak frequency of microbial infection using the methodology described above, with respect to the simulated systems' ability to attain given levels of haloacetic acids and total trihalomethanes. A more detailed set of results is presented in Appendix J. Results are presented for systems complying with SWTR and enhanced SWTR disinfection levels and consider the use of alternative disinfectants to chlorine. Endemic levels are expressed in number of cases per 10,000 per year, and outbreak frequency is expressed as the average number of years between outbreaks.

The results indicate that systems only minimally meeting SWTR standards

(3-log removal and inactivation of Giardia, 4-log removal and inactivation of viruses, and maintenance of a disinfectant residual in the distribution system) could produce water yielding significant endemic levels of microbial illness under different potential TTHM or HAA drinking water standards. Additionally, efforts to lower HAA or TTHM levels by reducing chlorine disinfection could not only increase endemic illnesses, but might lead to frequent outbreaks of illness in the community. The implications for systems in the upper percentiles of the distribution, which likely represents systems with poorer quality source waters, are particularly worrisome. The data indicate a precipitous decrease in the time between outbreaks as the distributions moves from the mean into the upper percentiles.

It is important to note that these results represent what might occur if all systems were only to minimally meet the SWTR requirements and treatment constraints as described by Gelderloos, et al (1992). Since many systems now provide higher levels of inactivation from the minimums used for this modeling analysis, the predicted infection rates and outbreak frequency rates may significantly overestimate what actually might occur, especially under the current TTHM MCL of 100 ug/l and corresponding high MCL target of 50-60 ug/l for HAAs.

However, as the DBP MCL targets decrease and it becomes more difficult for more systems to meet such a target, there should be greater likelihood for systems to only minimally meet the SWTR requirements. Therefore, greater significance should be given to the predicted infection and outbreak incidence at the lower DBP MCLs and to the difference between the predicted values under high versus low DBP MCL targets.

On the other hand, increasing the level of disinfection for poorer source waters (i.e., in accordance with EPA guidance to the SWTR) to an "enhanced" SWTR could reduce endemic illness and outbreaks to de minimus levels. The data from modeling show typically 1,000-fold lower numbers of infections per year.

An important question for any computer modeling effort is whether the simulation comes close to matching reality as well as we can estimate it. The statistical models used to describe the observed occurrence data of Hibler (1988) and LeChevallier, et al (1991) were seen to be poorest in their fit towards the high occurrence end of the distributions. This is critical, since the data from the model suggest that the majority of outbreaks occur from systems described by this part of the distribution. However, we believe that any overestimation of outbreaks from the model can be compensated for in the interpretation of the data.

The data from these simulations indicate a 2-5% annual risk of Giardia infection to the first customer drinking nominal SWTR-treated water. We expect that the increased disinfection CT farther in the distribution system should yield a much lower infection rate over the entire population served. For the population at large benefiting from this additional CT, we estimate that this infection rate would probably be reduced by an order of magnitude, i.e., the average infection rate would range from 0.2-0.5% per year for the population as a whole, if systems have only to meet the SWTR requirements.

For the 103 million people in the U.S. represented by our model simulation, this translates to an endemic level of about 200,000-500,000 infections from Giardia each year. The treatment system modeled here was chosen by Gelderloos, et al (1992) to represent conventional treatment systems using surface water most likely to produce drinking water capable of meeting SWTR standards. Treatment plants for the remaining 60 million people in the U.S. served by surface water are not believed to be as effective. We believe that these systems are likely to yield an additional 200,000-500,000 infections from Giardia each year, for a total U.S. infection rate of about 400,000-1 million per year. (Systems using groundwater exclusively as their source water contribute few, if any, Giardia infections.) If our assumptions on cyst viability and illness/infection rates are valid (see also Regli, et al, 1991), then perhaps 10% of these predicted infections will result in illness, or about 40,000-100,000 cases per year. Data from the Centers for Disease Control (Bennett, et al, 1987) indicate that Giardia contributes about 70,000 cases of illness in the U.S. each year, in very good agreement with our estimates.

The effectiveness of an enhanced SWTR disinfection is amplified by the consideration of all waterborne microbial illness. It is estimated that 940,000 cases of waterborne microbial illness occurred in 1985, resulting in some 900 deaths (Bennett, et al, 1987). (While Giardia is not considered to contribute to microbially-related deaths, the overall death rate from waterborne microbial illness is about 0.1%.) Giardia was chosen for our calculations in part due to its resistance to disinfection, which is generally greater than that for bacteria and viruses. Yet modeling of enhanced SWTR versus nominal SWTR treatment indicated an additional 3-log decrease in giardiasis by employing the enhanced SWTR. This could reasonably be expected to apply to disinfection of bacteria and viruses as well, which could result in substantial decreases in overall endemic microbial illness from drinking water and reduce related deaths to a de minimus level. Even if outbreak occurrence rates predicted from the model for minimal SWTR treatment are overestimated, it is clear that increased disinfection to the enhanced SWTR will eliminate treatment-system derived outbreaks.

REFERENCES

Bennett JV, SD Holmberg, MF Rogers and SL Solomon (1987). Infectious and papsitic diseases. Am J Prev Med 3: 102-114. In RW Amler and HB Dull (eds), Closing the gap: the burden of unnecessary illness. Oxford University Press, pp 102-114.

Beyer, WH, ed. (1968). CRC Handbook of Tables for Probability and Statistics. The Chemical Rubber Company, Cleveland.

Cromwell, JE, X Zhang, FJ Letkiewicz, S Regli and BA Macler (1992), Analysis of Potential Trade-offs in Regulation of Disinfection By-Products, U.S. Environmental Protection Agency.

Gelderloos, AB, et al (1992). Simulation of Compliance Choices for the Disinfection By-Products Regulatory Impact Analysis, Seminar on Disinfection and Disinfection By-Product Control, 1992 Annual Conference of the American Water Works Association, Vancouver, B.C.

Hibler CP (1988). Analysis of municipal water samples for cysts of Giardia. In: Wallis P and B Hammon (eds). Advances in Giardia research. Univ of Calgary Press, ppp 237-245.

Jolley LBW (1961). Summation of Series. Dover Publications, Inc., New York.

LeChevallier MW, WD Norton and RG Lee (1991). Occurrence of Giardia and Cryptosporidium spp. in surface water supplies. Appl Env Microb 57: 2610-2616.

Regli S, JB Rose, CN Haas and CP Gerba (1991). Modeling the risk from Giardia and viruses in drinking water. J Am Water Works Assoc 83:11: 76-84.

Rendtorff RC (1954). The experimental transmission of human intestinal protozoan parasites. II. Giardia lamblia cystes given in capsules. Am J Hyg 59: 209-220.

Rendtorff RC and CJ Holt (1954). The experimental transmission of human intestinal protozoan parasites. IV. Attempts to transmit Endamoeba coli and Giardia lamblia by water. Am J Hyg 60: 327-328.

Rose JB, CN Haas and S Regli (1991). Risk assessment and control of waterborne giardiasis. Am J Pub Health 81: 709-713.

Shapiro SS and MB Wilk (1965). An analysis of variance test for normality (complete samples). Biometrika 52: 591-611.

SAS Procedures Guide, version 6, third edition (1990). SAS Institute, Inc., Cary, North Carolina.

TABLE 1. SUMMARY OF ENDEMIC LEVELS AND OUTBREAK FREQUENCY FOR HALOACETIC ACID AND TOTAL TRIHALOMETHANES REGULATORY ALTERNATIVES

Rule	Disinfection	HAA	Average Number of Cases per 10,000 per Year	Mean Average Number of Years for an Outbreak	90th %ile Average Number of Years for an Outbreak
SWTR	With	60	210	∞	3,100
SWTR	With	50	230	∞	2.5
SWTR	With	40	280	2.5 E13	0.66
SWTR	With	30	320	9.3 E9	0.17
SWTR	With	20	400	5.3 E4	0.018
SWTR	With	10	560	4.0	0.006
SWTR	Without	60	240	∞	210
SWTR	Without	50	250	∞	1.1
SWTR	Without	40	270	∞	0.66
SWTR	Without	30	320	1.1 E10	0.17
SWTR	Without	20	390	1.3 E5	0.034
SWTR	Without	10	490	73	0.008
ESWTR	With	60	0.26	∞	∞
ESWTR	With	50	0.26	∞	∞
ESWTR	With	40	0.28	∞	∞
ESWTR	With	30	0.31	∞	∞
ESWTR	With	20	0.31	∞	∞
ESWTR	With	10	0.31	∞	∞
ESWTR	Without	60	0.25	∞	∞
ESWTR	Without	50	0.26	∞	∞
ESWTR	Without	40	0.27	∞	∞
ESWTR	Without	30	0.31	∞	∞
ESWTR	Without	20	0.32	∞	∞
ESWTR	Without	10	0.32	∞	∞

Assumptions and definitions:

- predicted incidence at first customer
- arithmetic average number of cysts across cities for average number of cases per 10,000 per year
- 90th percentile of distribution of number of cysts across cities for average number of years for an outbreak
- 25% secondary infection rate
- 1-log reduction in treatment performance 5% of the time
- SWTR = surface water treatment rule
- ESWTR = enhanced surface water treatment rule
- HAA = haloacetic acid target MCL (μg/l)

TABLE 1. (continued) SUMMARY OF ENDEMIC LEVELS AND OUTBREAK FREQUENCY FOR HALOACETIC ACID AND TOTAL TRIHALOMETHANES REGULATORY ALTERNATIVES

Rule	Alternative Disinfection	TTHM	Average Number of Cases per 10,000 per Year	Mean Average Number of Years for an Outbreak	90th %ile Average Number of Years for an Outbreak
SWTR	With	100	330	5.1 E8	3.0
SWTR	With	75	380	5.7 E5	0.18
SWTR	With	50	460	440	0.035
SWTR	With	25	500	53	0.008
SWTR	Without	100	340	1.9 E8	1.1
SWTR	Without	75	370	1.0 E6	0.18
SWTR	Without	50	460	380	0.035
SWTR	Without	25	500	58	0.008
ESWTR	With	100	0.26	∞	∞
ESWTR	With	75	0.29	∞	∞
ESWTR	With	50	0.30	∞	∞
ESWTR	With	25	0.32	∞	∞
ESWTR	Without	100	0.25	∞	∞
ESWTR	Without	75	0.28	∞	∞
ESWTR	Without	50	0.30	∞	∞
ESWTR	Without	25	0.32	∞	∞

Assumptions and definitions:

- predicted incidence at first customer
- arithmetic average number of cysts across cities for average number of cases per 10,000 per year
- 90th percentile of distribution of number of cysts across cities for average number of years for an outbreak
- 25% secondary infection rate
- 1-log reduction in treatment performance 5% of the time
- SWTR = surface water treatment rule
- ESWTR = enhanced surface water treatment rule
- TTHM = total trihalomethane level ($\mu g/l$)

APPENDIX A

LISTING OF RAW WATER DATA FROM LE CHEVALLIER USED TO CHARACTERIZE
GEOGRAPHIC VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.08 TO REFLECT RETRIEVAL EFFICIENCY
MEASUREMENTS HAVE ALSO BEEN MULTIPLIED BY 0.13 AS A VIABILITY FACTOR

CITY	ANNUAL AVERAGE NO. OF CYSTS IN RAW WATER	OBSERVATION WAS DELETED (*)
1	506.73	
2	1758.79	
101	184.44	*
102	22.48	
109	86.40	
109	49.13	*
109	75.16	*
302	16.25	
305	720.42	
306	4.33	
307	3087.50	
307	174.15	
307	1084.42	
307	2143.38	
307	1760.42	
307	1760.42	
307	1706.25	
307	1706.25	
310	151.67	
310	61.48	
311	189.31	
312	102.65	
312	894.02	
312	894.02	
314	80.71	*
315	22.56	*
401	54.17	*
402	22.48	
404	121.87	
405	12.59	
406	4.93	
407	110.77	
409	3096.98	
410	1407.79	
410	135.42	
411	182.81	
414	270.83	
501	6770.83	
502	39.54	*
502	366.55	*
502	157.76	*
503	331.77	
504	1692.71	
504	108.33	
504	13.54	
504	595.83	
504	54.17	
506	1780.73	

LISTING OF RAW WATER DATA FROM LE CHEVALLIER USED TO CHARACTERIZE
GEOGRAPHIC VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.08 TO REFLECT RETRIEVAL EFFICIENCY
MEASUREMENTS HAVE ALSO BEEN MULTIPLIED BY 0.13 AS A VIABILITY FACTOR

CITY	ANNUAL AVERAGE NO. OF CYSTS IN RAW WATER	OBSERVATION WAS DELETED (*)
508	32.50	
509	108.33	
511	89.92	
512	195.00	
512	514.58	
513	1468.73	
514	37.92	
516	410.85	
517	124.96	
518	162.50	
519	541.67	
602	338.54	
603	90.46	
604	2499.25	
605	601.79	
605	622.37	
605	855.02	
605	270.83	
605	542.21	
605	1303.52	
605	277.60	
605	1516.67	
606	542.21	
606	277.60	
608	376.19	*
609	180.65	*
610	492.37	
611	29.01	*
612	801.94	
613	1269.53	
614	297.92	
615	24808.33	*
616	98.58	
618	108.87	
619	60.94	
703	52.00	*
703	19.34	*

LISTING OF RAW WATER DATA FROM LE CHEVALLIER USED TO CHARACTERIZE
GEOGRAPHIC VARIATION (IN CYSTS PER 100 GALLONS)
MULTIPLE MEASUREMENTS AT A CITY AVERAGED
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.08 TO REFLECT RETRIEVAL EFFICIENCY
MEASUREMENTS HAVE ALSO BEEN MULTIPLIED BY 0.13 AS A VIABILITY FACTOR

CITY	NUMBER OF MEASUREMENTS	CITY AVERAGE
1	1	506.73
2	1	1758.79
102	1	22.48
109	1	86.40
302	1	16.25
305	1	720.42
306	1	4.33
307	8	1677.85
310	2	106.57
311	1	189.31
312	3	630.23
402	1	22.48
404	1	121.87
405	1	12.59
406	1	4.93
407	1	110.77
409	1	3096.98
410	2	771.60
411	1	182.81
414	1	270.83
501	1	6770.83
503	1	331.77
504	5	492.92
506	1	1780.73
508	1	32.50
509	1	108.33
511	1	89.92
512	2	354.79
513	1	1468.73
514	1	37.92
516	1	410.85
517	1	124.96
518	1	162.50
519	1	541.67
602	1	338.54
603	1	90.46
604	1	2499.25
605	8	748.75
606	2	409.91
610	1	492.37
612	1	801.94
613	1	1269.53
614	1	297.92
616	1	98.58
618	1	108.87
619	1	60.94

SUMMARIES OF GIARDIA AND NATURAL LOGARITHM OF GIARDIA IN RAW WATER

UNIVARIATE PROCEDURE

Variable=GIARRAW GIARDIA IN RAW WATER

Moments

N	46	Sum Wgts	46
Mean	657.3852	Sum	30239.72
Std Dev	1149.549	Variance	1321462
Skewness	3.812403	Kurtosis	17.89069
USS	79344952	CSS	59465807
CV	174.8668	Std Mean	169.4917
T:Mean=0	3.878568	Prob>|T|	0.0003
Sgn Rank	540.5	Prob>|S|	0.0001
Num ^= 0	46		
W:Normal	0.572595	Prob<W	0.0001

Quantiles(Def=5)

100% Max	6770.833	99%	6770.833
75% Q3	720.4167	95%	2499.25
50% Med	284.375	90%	1758.792
25% Q1	90.45833	10%	22.47917
0% Min	4.333333	5%	12.59375
		1%	4.333333
Range	6766.5		
Q3-Q1	629.9583		
Mode	22.47917		

Extremes

Lowest	Obs	Highest	Obs
4.333333(	7)	1758.792(	2)
4.929167(	15)	1780.729(	24)
12.59375(	14)	2499.25(	37)
16.25(	5)	3096.979(	17)
22.47917(	12)	6770.833(	21)

UNIVARIATE PROCEDURE

Variable=LGGIAR NATURAL LOG OF GIARDIA IN RAW WATER

Moments

N	46	Sum Wgts	46
Mean	5.392429	Sum	248.0518
Std Dev	1.668332	Variance	2.783333
Skewness	-0.37578	Kurtosis	-0.08285
USS	1462.852	CSS	125.25
CV	30.93842	Std Mean	0.245982
T:Mean=0	21.92203	Prob>|T|	0.0001
Sgn Rank	540.5	Prob>|S|	0.0001
Num ^= 0	46		
W:Normal	0.975227	Prob<W	0.5824

Quantiles(Def=5)

100% Max	8.820379	99%	8.820379
75% Q3	6.57983	95%	7.823746
50% Med	5.649159	90%	7.472382
25% Q1	4.504889	10%	3.112589
0% Min	1.466337	5%	2.533201
		1%	1.466337
Range	7.354042		
Q3-Q1	2.07494		
Mode	3.112589		

Extremes

Lowest	Obs	Highest	Obs
1.466337(	7)	7.472382(	2)
1.59517(	15)	7.484778(	24)
2.533201(	14)	7.823746(	37)
2.788093(	5)	8.038182(	17)
3.112589(	12)	8.820379(	21)

APPENDIX B

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
1	0	9
	4	1
	6	2
	12	1
	24	1
	30	2
8	0	9
	2	1
	3	1
	5	2
	8	2
	12	1
	34	1
	63	1
	65	1
11	0	21
	3	1
	24	1
20	0	49
	1	1
	2	2
	4	4
	5	1
	6	2
	8	2
	28	1
32	0	19
	1	1
	6	1
	7	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
40	0	35
	1	2
	2	6
	3	1
	4	5
	6	1
	7	2
	8	2
	9	3
	10	1
	11	2
	12	3
	13	2
	14	1
	15	2
	17	1
	30	1
41	0	29
	2	4
	5	1
	6	1
	12	1
	13	1
	20	1
	48	1
	50	1
	51	1
61	0	28
	1	4
	2	3
	6	2
	10	2
	12	2
	13	1
	22	1
	40	1
	50	1
	162	1
62	0	2
	13	1
	48	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
64	0	24
	3	1
	4	1
	5	2
	6	2
69	6	1
	13	1
71	0	5
	1	1
	5	1
	20	1
73	0	3
	1	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
74	0	15
	2	7
	3	2
	4	2
	5	1
	6	2
	9	2
	11	4
	12	1
	13	1
	15	2
	20	1
	22	1
	29	1
	34	1
	36	1
	37	1
	38	1
	58	1
	66	1
	67	1
	95	1
	105	1
	130	1
	134	1
	147	1
	160	1
	204	1
	220	1
	252	1
	740	1
86	0	6
	1	2
	6	1
	9	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
89	0	90
	2	1
	3	4
	4	1
	5	5
	6	2
	8	1
	10	2
	11	2
	12	2
	13	1
	14	1
	17	1
	21	1
	24	1
	38	1
	42	1
90	0	184
	1	3
	2	2
	3	1
	5	3
	7	2
	9	1
	10	3
	11	1
	13	1
	16	1
	18	1
	23	1
	35	1
	52	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
91	0	170
	1	2
	2	2
	3	3
	4	1
	6	3
	7	1
	8	1
	9	2
	10	2
	11	2
	14	1
	17	2
	18	1
	30	1
	88	1
93	0	59
	2	1
	10	1
	13	1
	48	1
	50	1
	58	1
	59	1
	63	2
	88	1
	149	1
	191	1
	406	1
	433	1
	441	1
	448	1
	800	1
	836	1
	1301	1
100	0	7
	32	1
	45	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
101	0	2
	40	1
	200	1
	352	1
102	0	1
	10	1
105	0	4
	102	1
107	2	1
	16	1
109	0	52
	1	2
	2	3
	3	1
	4	2
	5	3
	6	1
	7	1
	8	1
	9	1
	11	1
	13	1
	16	1
	17	1
	24	1
111	0	4
	2	1
118	9	1
	93	1
120	0	52
	1	1
	2	1
121	0	16
	2	1
	5	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
122	0	2
	3	1
	243	1
130	0	9
	2	1
	3	2
	15	1
	17	1
	19	1
	22	1
136	6	1
139	0	4
142	0	16
	10	1
147	0	2
	2	1
	11	1
	16	1
	20	1
	28	1
	30	1
149	0	1
	27	1
	37	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
150	0	49
	1	5
	2	4
	3	1
	5	1
	6	2
	7	2
	8	1
	9	1
	11	1
	13	1
	14	1
	16	1
	17	2
	22	1
	24	2
	28	1
	29	1
	37	1
	40	1
	133	1
	146	1
152	0	27
	1	1
	2	1
	4	2
	5	1
	6	2
	7	1
	8	1
	9	1
	15	1
	16	1
	20	1
	23	1
	24	3
	26	1
	28	1
	31	3
	36	1
	106	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
159	0	2
	16	1
	27	1
160	2	1
162	0	7
	2	1
	6	1
	16	1
	19	2
164	2	1
165	0	3
	47	1
168	0	3
	2	1
	5	1
	8	1
	30	1
169	12	1
173	542	1
174	0	8
	12	1
	43	1
	46	1
	47	1
	55	1
175	0	4
	3	1
	4	1
	31	1
195	50	1
	54	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
199	0	34
	2	2
	4	1
	6	1
	10	1
	13	1
	17	1
201	18	1
207	0	1
	10	1
	14	1
208	0	1
	6	1
	7	1
212	6	1
215	0	1
	14	1
217	0	2
	2	1
	6	1
	11	1
229	20	1
230	17	1
	19	1
234	18	1
236	17	1
241	3	1
246	6	1
248	648	1
256	26	1

SUMMARY OF RAW WATER DATA FROM HIBLER USED TO CHARACTERIZE
TEMPORAL VARIATION (IN CYSTS PER 100 GALLONS)
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYSTS	FREQUENCY
257	86	1
258	4	1
259	0	1
	5	1
268	27	1
269	68	1
295	0	7
	2	1
	12	1
299	10	1
302	0	1
	8	1
308	0	7
	16	1

SUMMARY STATISTICS FOR RAW WATER DATA FROM HIBLER

MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYST MEASUREMENTS	PERCENTAGE OF MEAS. SHOWING ZERO CYSTS	MINIMUM CYST MEASUREMENT	MAXIMUM CYST MEASUREMENT	MEDIAN CYST MEASUREMENT	MEAN OF CYST MEASUREMENTS	STANDARD DEVIATION OF CYST MEASUREMENTS
7	16	56.250	0	30	0.0	7.000	11.027
8	19	47.368	0	65	2.0	10.789	20.376
11	23	91.304	0	24	0.0	1.174	5.015
20	62	79.032	0	28	0.0	1.323	3.995
32	22	86.364	0	7	0.0	0.636	1.916
40	70	50.000	0	30	0.5	4.071	5.906
41	41	70.732	0	51	0.0	5.195	13.312
61	46	60.870	0	162	0.0	7.674	25.326
62	4	50.000	0	48	6.5	15.250	22.677
64	30	80.000	0	6	0.0	0.967	2.025
69	2	0.000	6	13	9.5	9.500	4.950
71	8	62.500	0	20	0.0	3.250	6.985
73	4	75.000	0	1	0.0	0.250	0.500
74	59	25.424	0	740	9.0	46.712	109.488
86	10	60.000	0	9	0.0	1.700	3.164
89	117	76.923	0	42	0.0	2.547	6.674
90	206	89.320	0	52	0.0	1.194	5.230
91	195	87.179	0	88	0.0	1.518	7.241
93	78	75.641	0	1301	0.0	69.987	212.689
100	9	77.778	0	45	0.0	8.556	17.285
101	5	40.000	0	352	40.0	118.400	154.444
102	2	50.000	0	10	5.0	5.000	7.071
105	5	80.000	0	102	0.0	20.400	45.616
107	2	0.000	2	16	9.0	9.000	9.899
109	72	72.222	0	24	0.0	2.014	4.564
111	5	80.000	0	2	0.0	0.400	0.894
118	2	0.000	9	93	51.0	51.000	59.397
120	54	96.296	0	2	0.0	0.056	0.302
121	18	88.889	0	5	0.0	0.389	1.243
122	4	50.000	0	243	1.5	61.500	121.008
130	16	56.250	0	22	0.0	5.062	8.045
136	1	0.000	6	6	6.0	6.000	.
139	4	100.000	0	0	0.0	0.000	0.000
142	17	94.118	0	10	0.0	0.588	2.425
147	8	25.000	0	30	13.5	13.375	12.153
149	3	33.333	0	37	27.0	21.333	19.140
150	81	60.494	0	146	0.0	7.840	22.898
152	51	52.941	0	106	0.0	9.549	17.715
159	4	50.000	0	27	8.0	10.750	13.200
160	1	0.000	2	2	2.0	2.000	.
162	12	58.333	0	19	0.0	5.167	7.964
164	1	0.000	2	2	2.0	2.000	.
165	4	75.000	0	47	0.0	11.750	23.500
168	7	42.857	0	30	2.0	6.429	10.830
169	1	0.000	12	12	12.0	12.000	.
173	1	0.000	542	542	542.0	542.000	.
174	13	61.538	0	55	0.0	15.615	22.681
175	7	57.143	0	31	0.0	5.429	11.400
195	2	0.000	50	54	52.0	52.000	2.828
199	41	82.927	0	17	0.0	1.317	3.698

SUMMARY STATISTICS FOR RAW WATER DATA FROM HIBLER
MEASUREMENTS HAVE BEEN MULTIPLIED BY 2.0 TO REFLECT RETRIEVAL EFFICIENCY AND ROUNDED

CITY	NUMBER OF CYST MEASUREMENTS	PERCENTAGE OF MEAS. SHOWING ZERO CYSTS	MINIMUM CYST MEASUREMENT	MAXIMUM CYST MEASUREMENT	MEDIAN CYST MEASUREMENT	MEAN OF CYST MEASUREMENTS	STANDARD DEVIATION OF CYST MEASUREMENTS
201	1	0.0000	18	18	18.0	18.000	.
207	3	33.3333	0	14	10.0	8.000	7.21110
208	3	33.3333	0	7	6.0	4.333	3.78594
212	1	0.0000	6	6	6.0	6.000	.
215	2	50.0000	0	14	7.0	7.000	9.89949
217	5	40.0000	0	11	2.0	3.800	4.71169
229	1	0.0000	20	20	20.0	20.000	.
230	2	0.0000	17	19	18.0	18.000	1.41421
234	1	0.0000	18	18	18.0	18.000	.
236	1	0.0000	17	17	17.0	17.000	.
241	1	0.0000	3	3	3.0	3.000	.
246	1	0.0000	6	6	6.0	6.000	.
248	1	0.0000	648	648	648.0	648.000	.
256	1	0.0000	26	26	26.0	26.000	.
257	1	0.0000	86	86	86.0	86.000	.
258	1	0.0000	4	4	4.0	4.000	.
259	2	50.0000	0	5	2.5	2.500	3.53553
268	1	0.0000	27	27	27.0	27.000	.
269	1	0.0000	68	68	68.0	68.000	.
295	9	77.7778	0	12	0.0	1.556	3.97213
299	1	0.0000	10	10	10.0	10.000	.
302	2	50.0000	0	8	4.0	4.000	5.65685
308	8	87.5000	0	16	0.0	2.000	5.65685

APPENDIX C

Expected Value [E(X)] of the Delta-Negative Binomial Distribution

$$\sum_{x=0}^{\infty} x\left[\delta \bullet I_0(x)+(1-\delta)\bullet\frac{\Gamma(q+x)}{\Gamma(q)\bullet x!}\bullet\left(\frac{q}{q+m}\right)^{q}\bullet\left(\frac{m}{q+m}\right)^{x}\right]$$

$$=(1-\delta)\sum_{x=0}^{\infty} x\bullet\frac{\Gamma(q+x)}{\Gamma(q)\bullet x!}\bullet\left(\frac{q}{q+m}\right)^{q}\bullet\left(\frac{m}{q+m}\right)^{x}$$

$$=(1-\delta)\bullet\left(\frac{q}{q+m}\right)^{q}\sum_{x=0}^{\infty} x\bullet\frac{\Gamma(q+x)}{\Gamma(q)\bullet x!}\bullet\left(\frac{m}{q+m}\right)^{x}$$

$$=(1-\delta)\bullet\left(\frac{q}{q+m}\right)^{q}\sum_{x=1}^{\infty}\frac{\Gamma(q+x)}{\Gamma(q)\bullet(x-1)!}\bullet\left(\frac{m}{q+m}\right)^{x}$$

Let j = x-1

$$=(1-\delta)\bullet\left(\frac{q}{q+m}\right)^{q}\bullet\left(\frac{m}{q+m}\right)\sum_{j=0}^{\infty}\frac{\Gamma(q+j+1)}{\Gamma(q)\bullet j!}\bullet\left(\frac{m}{q+m}\right)^{j}$$

$$=(1-\delta)\bullet\left(\frac{q}{q+m}\right)^{q}\bullet\left(\frac{m}{q+m}\right)\sum_{j=0}^{\infty}\frac{(q+j)\bullet\Gamma(q+j)}{\Gamma(q)\bullet j!}\bullet\left(\frac{m}{q+m}\right)^{j},\ \textit{since}\ \Gamma(a+1)=a\bullet\Gamma(a)$$

$$=(1-\delta)\bullet q\bullet\left(\frac{q}{q+m}\right)^{q}\bullet\left(\frac{m}{q+m}\right)\bullet\left(\frac{q}{q+m}\right)^{-(q+1)}\qquad \textit{(see note below)}$$

$$=(1-\delta)\bullet q\bullet\frac{m}{q}$$

$$=(1-\delta)\bullet m$$

From Jolley (1961), formula 1015:

$$\sum_{j=0}^{\infty}\frac{(q+j)\bullet\Gamma(q+j)}{\Gamma(q)\bullet j!}\bullet\left(\frac{m}{q+m}\right)^{j}=q\,[1+(q+1)\bullet\left(\frac{m}{q+m}\right)+$$
$$\frac{(q+1)\bullet(q+2)}{2!}\bullet\left(\frac{m}{q+m}\right)^{2}+\ldots]=q\left[1-\left(\frac{m}{q+m}\right)\right]^{-(q+1)}=q\bullet\left(\frac{q}{q+m}\right)^{-(q+1)}$$

Variance [V(X)] of the Delta-Negative Binomial Distribution

$$V(X) = E(X^2) - [E(X)]^2$$
$$= E(X^2) - [(1-\delta)\cdot m]^2$$

$$=\sum_{x=0}^{\infty} x^2\left[\delta\cdot I_0(x) + (1-\delta)\cdot\frac{\Gamma(q+x)}{\Gamma(q)\cdot x!}\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^x\right] - [(1-\delta)\cdot m]^2$$

$$=\sum_{x=2}^{\infty} x\cdot(x-1)\cdot(1-\delta)\cdot\frac{\Gamma(q+x)}{\Gamma(q)\cdot x!}\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^x + (1-\delta)\cdot m - [(1-\delta)^2\cdot m^2] \quad \textit{(see note below)}$$

$$=\sum_{x=2}^{\infty} (1-\delta)\cdot\frac{\Gamma(q+x)}{\Gamma(q)\cdot(x-2)!}\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^x + (1-\delta)\cdot m - [(1-\delta)^2\cdot m^2]$$

Let j=x-2

$$=(1-\delta)\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^2\sum_{j=0}^{\infty}\frac{\Gamma(q+j+2)}{\Gamma(q)\cdot j!}\cdot\left(\frac{m}{q+m}\right)^j + (1-\delta)\cdot m - [(1-\delta)^2\cdot m^2]$$

$$=(1-\delta)\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^2\sum_{j=0}^{\infty}\frac{(q+j)\cdot(q+j+1)\cdot\Gamma(q+j)}{\Gamma(q)\cdot j!}\cdot\left(\frac{m}{q+m}\right)^j + (1-\delta)\cdot m - [(1-\delta)^2\cdot m^2]$$

[*since* $\Gamma(a+2)=(a+1)\cdot a\cdot\Gamma(a)$]

$$=(1-\delta)\cdot\left(\frac{q}{q+m}\right)^q\cdot\left(\frac{m}{q+m}\right)^2\cdot(q+1)\cdot q\cdot\left(\frac{q}{q+m}\right)^{-(q+2)} + (1-\delta)\cdot m - [(1-\delta)^2\cdot m^2] \quad \textit{(see note below)}$$

$$=(1-\delta)\cdot(q+1)\cdot q\cdot\frac{m^2}{q^2} + (1-\delta)\cdot m - [(1-\delta)^2\cdot m^2]$$

$$=(1-\delta)\cdot m\left[m+1+\frac{m}{q}-(1-\delta)\cdot m\right]$$

$$=(1-\delta)\cdot m\left[(\delta\cdot m)+1+\frac{m}{q}\right]$$

Notes: $E(X^2) = E[X\cdot(X-1)] + E(X)$.
Also, from Jolley (1961), formula 1015:

$$\sum_{j=0}^{\infty}\frac{(q+j)\cdot(q+j+1)\cdot\Gamma(q+j)}{\Gamma(q)\cdot j!}\cdot\left(\frac{m}{q+m}\right)^j = q\cdot(q+1)\cdot\left[1+(q+2)\cdot\left(\frac{m}{q+m}\right)+\right.$$
$$\left.\frac{(q+2)\cdot(q+3)}{2!}\cdot\left(\frac{m}{q+m}\right)^2+\ldots\right] = q\cdot(q+1)\left[1-\left(\frac{m}{q+m}\right)\right]^{-(q+2)} = q\cdot(q+1)\cdot\left(\frac{q}{q+m}\right)^{-(q+2)}.$$

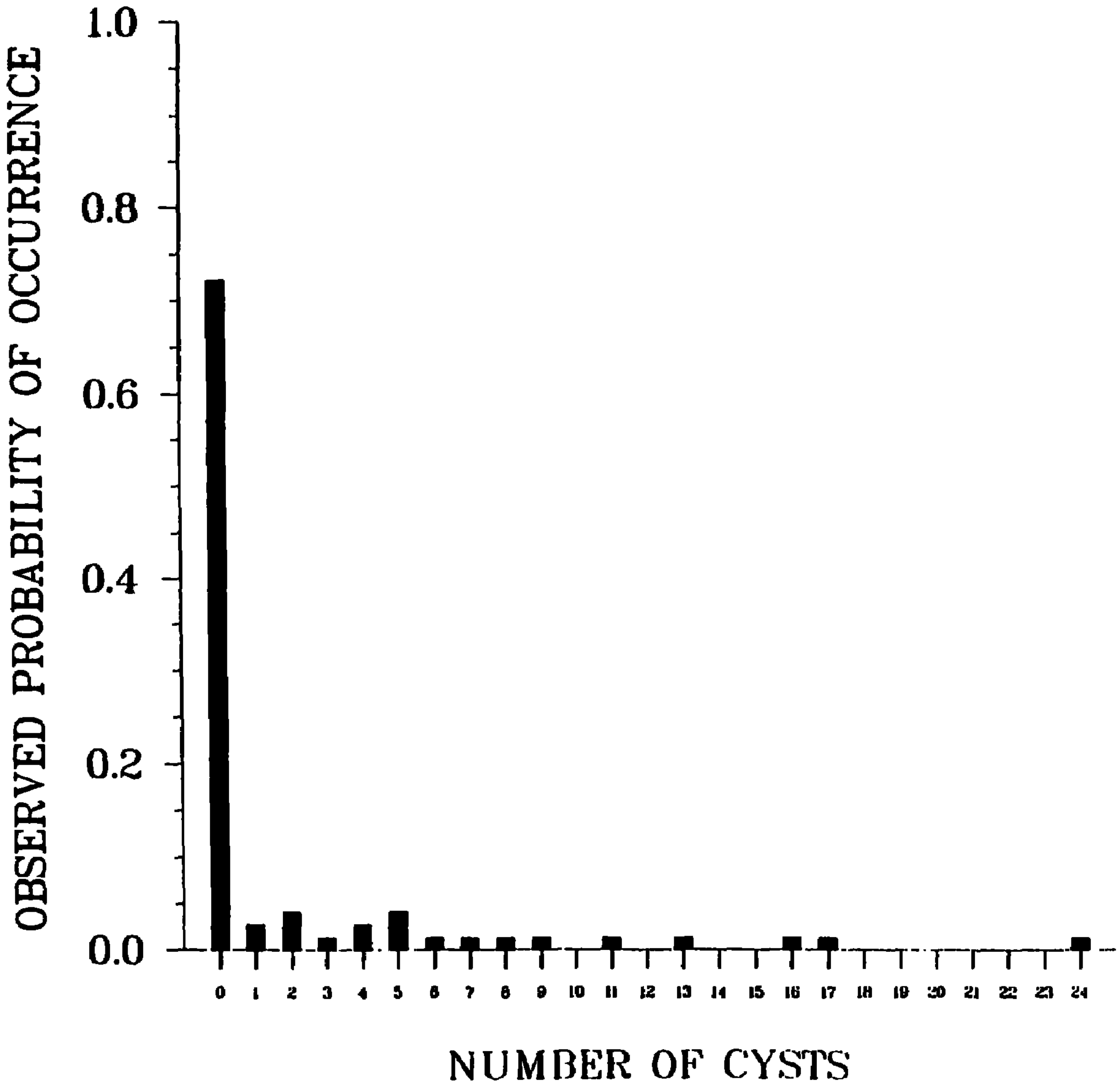
APPENDIX D

City	Sample Size	q	m	δ
90	206	0.717917	9.527345	0.874538
91	195	0.677301	10.028699	0.848751
89	117	1.625189	10.619969	0.760072
150	81	0.209461	11.264430	0.304341
93	78	0.553427	278.938986	0.749330
109	72	1.341256	6.579215	0.693713
40	70	1.874561	7.781179	0.477132
20	62	1.301950	5.592729	0.758171
74	59	0.256774	46.231592	0
120*	54	∞	1.500000	0.962963
152	51	1.102640	19.489594	0.510031
61	46	0.112547	6.507630	0
41	41	0.441110	13.936982	0.628038
199	41	1.953953	7.351698	0.821713
64	30	∞	4.833013	0.800007
11*	23	0.880435	13.500000	0.913043
32	22	2.707007	4.302894	0.849712
8	19	0.570479	17.162917	0.381082
121	18	∞	3.519917	0.885147
142*	17	∞	10.000000	0.941176
1	16	2.038864	15.809209	0.557102
130	16	1.355463	10.995573	0.543141
174	13	5.520679	40.699701	0.626873
162	12	2.272272	12.213228	0.577371
86	10	0.254426	1.811588	0
100	9	376.89806	38.499536	0.775875
295	9	1.446761	6.684823	0.765451
71	8	0.294975	5.006120	0.358253
147	8	2.346651	17.580409	0.251390
308	8	∞	15.999823	0.874989
168	7	0.914368	10.568594	0.420023
175	7	0.621825	10.508781	0.487783
105*	5	∞	102.000000	0.800000
101	5	1.653133	198.003206	0.402287
111	5	∞	1.633307	0.752298
217	5	3.696419	6.163064	0.383741
62	4	2.932655	30.968027	0.510882
73	4	∞	0.326039	0.222424
122	4	0.281353	117.737349	0.488077
139*	4	∞	0.000000	1.000000
159	4	50.514247	21.498798	0.515855
165	4	∞	46.979735	0.750057

Note: * denotes starting values for these parameters. Nonlinear algorithm failed to detect any optimum parameter estimates away from the starting values.

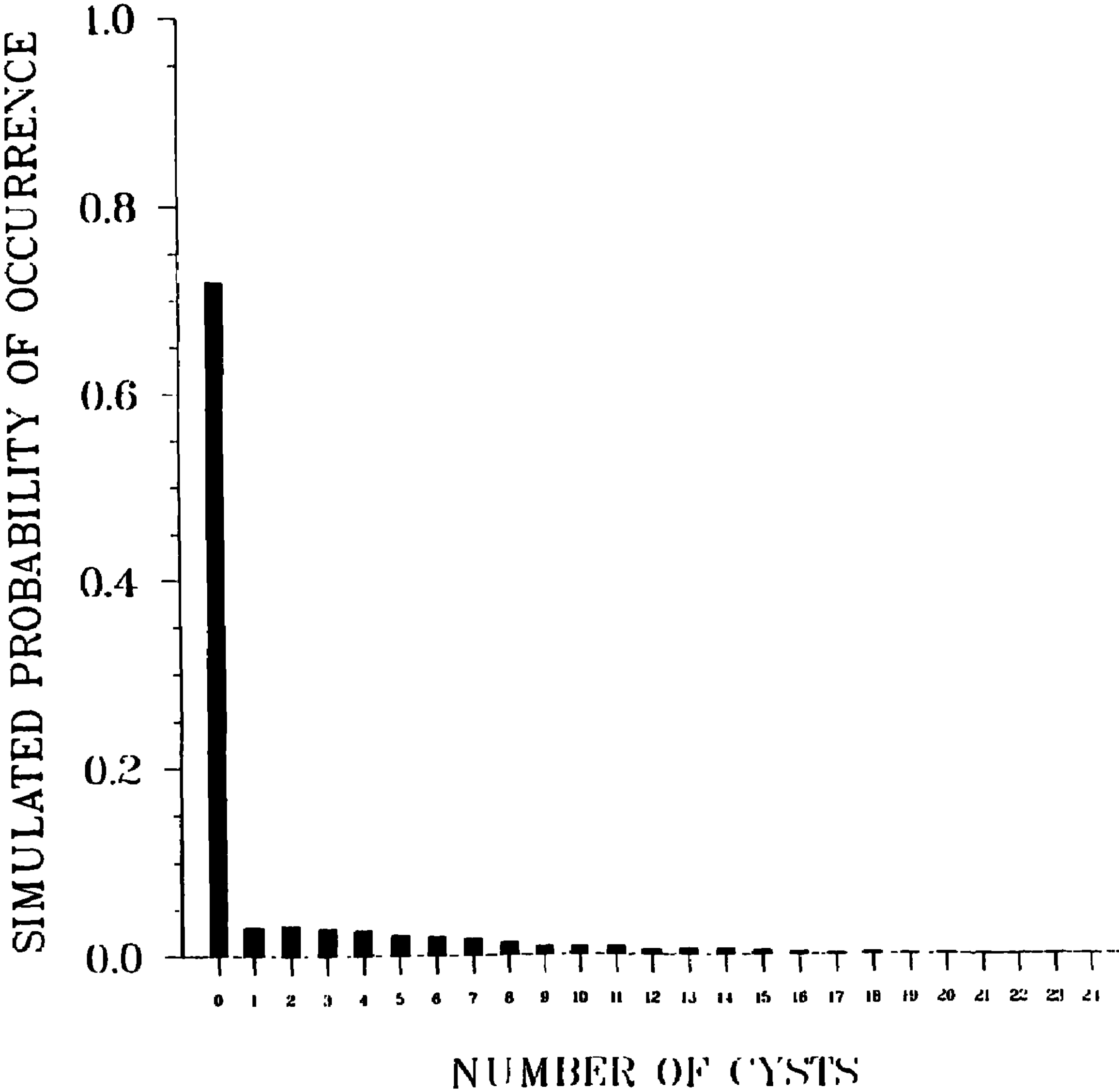
APPENDIX E

ACTUAL CITY 109 GIARDIA OBSERVATIONS .
(N=72)



APPENDIX F

SIMULATION BASED ON THE DELTA-NEGATIVE BINOMIAL MODEL PARAMETERS FOR CITY 109 (N=10,000)



APPENDIX G

Application of a Binomial Removal Process to a Delta-Negative Binomial Distribution

Let X = number of cysts in raw water
and K = number of cysts in finished water.

For k=0 and x=0, Pr(K=0)=δ.

For k≥0 and x>0 (x≥k):

$$Pr(K=k) = (1-\delta)\sum_{x=k}^{\infty}\frac{x!}{k!\cdot(x-k)!}\cdot p^{k}\cdot(1-p)^{x-k}\cdot\frac{\Gamma(q+x)}{\Gamma(q)\cdot x!}\cdot\left(\frac{q}{q+m}\right)^{q}\cdot\left(\frac{m}{q+m}\right)^{x}$$

$$=(1-\delta)\cdot\frac{p^{k}}{k!\cdot\Gamma(q)}\cdot\left(\frac{q}{q+m}\right)^{q}\left[\sum_{x=k}^{\infty}\frac{1}{(x-k)!}\cdot(1-p)^{x-k}\cdot\Gamma(q+x)\cdot\left(\frac{m}{q+m}\right)^{x}\right]$$

Let j = x-k

$$=(1-\delta)\cdot\frac{p^{k}}{k!\cdot\Gamma(q)}\cdot\left(\frac{q}{q+m}\right)^{q}\cdot\left(\frac{m}{q+m}\right)^{k}\left[\sum_{j=0}^{\infty}\frac{1}{j!}\cdot(1-p)^{j}\cdot\Gamma(q+k+j)\cdot\left(\frac{m}{q+m}\right)^{j}\right]$$

$$=(1-\delta)\cdot p^{k}\cdot\left(\frac{q}{q+m}\right)^{q}\cdot\left(\frac{m}{q+m}\right)^{k}\cdot\frac{\Gamma(q+k)}{k!\cdot\Gamma(q)}\left[\sum_{j=0}^{\infty}\frac{\Gamma(q+k+j)}{\Gamma(q+k)\cdot j!}\cdot\left(\frac{(1-p)\cdot m}{q+m}\right)^{j}\right]$$

$$=(1-\delta)\cdot p^{k}\cdot\left(\frac{q}{q+m}\right)^{q}\cdot\left(\frac{m}{q+m}\right)^{k}\cdot\frac{\Gamma(q+k)}{k!\cdot\Gamma(q)}\left[1-\left(\frac{(1-p)\cdot m}{q+m}\right)\right]^{-(q+k)}$$ *(see note below)*

$$=(1-\delta)\cdot\left(\frac{q}{q+m}\right)^{q}\cdot\left(\frac{m\cdot p}{q+m}\right)^{k}\cdot\frac{\Gamma(q+k)}{k!\cdot\Gamma(q)}\cdot\left[\frac{q+(m\cdot p)}{q+m}\right]^{-(q+k)}$$

$$=(1-\delta)\cdot\frac{\Gamma(q+k)}{k!\cdot\Gamma(q)}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{(m\cdot p)}{q+(m\cdot p)}\right)^{k}.$$

Consequently, $K \sim DNB[\delta,(m\cdot p),q]$.

From Jolley, 1961, formula 1015:

$$\sum_{x=0}^{\infty}\frac{\Gamma(q+k+j)}{\Gamma(q+k)\cdot j!}\cdot\left(\frac{(1-p)\cdot m}{q+m}\right)^{j}=1+(q+k)\cdot\left(\frac{(1-p)\cdot m}{q+m}\right)+$$

$$\frac{(q+k)\cdot(q+k+1)}{2!}\cdot\left(\frac{(1-p)\cdot m}{q+m}\right)^{2}+\ldots=\left[1-\left(\frac{(1-p)\cdot m}{q+m}\right)\right]^{-(q+k)}=\left(\frac{q+(m\cdot p)}{q+m}\right)^{-(q+k)}.$$

APPENDIX H

Non-Linear Least Squares Iterative Phase Dependent Variable RESP Method: Gauss-Newton

Iter	R	Sum of Squares
0	0	21.000000
1	0.0000014043	16.593611
2	0.0000043460	14.226365
3	0.0000110043	12.735403
4	0.0000259919	11.781231
5	0.0000837003	10.500940
6	0.000190	9.928268
7	0.000343	9.755801
8	0.001451	9.036939
9	0.012521	5.940740
10	0.017625	5.729831
11	0.017917	5.729345
12	0.017907	5.729345
13	0.017907	5.729345

NOTE: Convergence criterion met.

Non-Linear Least Squares Summary Statistics Dependent Variable RESP

Source	DF	Sum of Squares	Mean Square
Regression	1	15.270655199	15.270655199
Residual	39	5.729344801	0.146906277
Uncorrected Total	40	21.000000000	
(Corrected Total)	39	9.975000000	

Parameter	Estimate	Asymptotic Std. Error	Asymptotic 95 % Confidence Interval Lower	Upper
R	0.0179074539	0.00502972826	0.00773391748	0.02808099039

Asymptotic Correlation Matrix

Corr	R
R	1

APPENDIX I

Variance of the Exponential Dose-Response Function where Dose (K) Has a Delta-Negative Binomial Distribution (continued)

Now, $E(X^2) = E(X^2) - [E(X)]^2$

$$So\cdot E(X^2) = (1-\delta)\left[1 - 2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q + \left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-2\cdot r^*}))}\right)^q\right]$$
$$- (1-\delta)^2\left[1 - \left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q\right]^2$$

$$= (1-\delta)\left[1 - 2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q + \left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-2\cdot r^*}))}\right)^q\right]$$
$$- (1-\delta)^2\left[1 - 2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q + \left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^{2\cdot q}\right]$$

$$= (1-\delta)\cdot[1-(1-\delta)]\cdot\left[1 - 2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q\right]$$
$$+ (1-\delta)\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-2\cdot r^*}))}\right)^q - (1-\delta)^2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^{2\cdot q}$$

$$= (1-\delta)\left\{\delta\left[1 - 2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q\right]\right.$$
$$\left.+ \left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-2\cdot r^*}))}\right)^q - (1-\delta)\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^{2\cdot q}\right\}$$

Expected Value of the Exponential Dose-Response Function Where Dose (K) Has a Delta-Negative Binomial Distribution

$$E[1-e^{(-r^{*}\cdot K)}]=\sum_{k=0}^{\infty}\left[1-e^{(-r^{*}\cdot k)}\right]\left[\delta\cdot I_{0}(k)+(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^{k}\right]$$

$$=\delta\cdot(1-\exp(0))+\sum_{k=0}^{\infty}\left[1-e^{(-r^{*}\cdot k)}\right]\left[(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^{k}\right]$$

$$=\sum_{k=0}^{\infty}\left[1-e^{(-r^{*}\cdot k)}\right]\left[(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^{k}\right]$$

$$=(1-\delta)\sum_{k=0}^{\infty}\left[\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\right]\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^{k}-(1-\delta)\sum_{k=0}^{\infty}e^{(-r^{*}\cdot k)}\left[\frac{\Gamma(q+x)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{m\cdot p)}{q+(m\cdot p)}\right)^{k}\right]$$

$$=(1-\delta)\ -\ (1-\delta)\sum_{k=0}^{\infty}e^{(-r^{*}\cdot k)}\left[\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^{k}\right]$$

$$=(1-\delta)\ -\ (1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\left[\sum_{k=0}^{\infty}\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{m\cdot p\cdot e^{-r^{*}}}{q+(m\cdot p)}\right)^{k}\right]$$

$$=(1-\delta)\ -\ (1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\left[1-\left(\frac{m\cdot p\cdot e^{-r^{*}}}{q+(m\cdot p)}\right)\right]^{-q}. \qquad \textit{(see note below)}$$

$$=(1-\delta)\ -\ (1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)}\right)^{q}\cdot\left(\frac{q+(m\cdot p)-(m\cdot p\cdot e^{-r^{*}})}{q+(m\cdot p)}\right)^{-q}$$

$$=(1-\delta)\ -\ (1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)-(m\cdot p\cdot e^{-r^{*}})}\right)^{q}$$

$$=(1-\delta)\left\{1-\left[\frac{q}{q+(m\cdot p\cdot(1-e^{-r^{*}}))}\right]^{q}\right\}$$

From Jolley, 1961, formula 1015:

$$\sum_{k=0}^{\infty}\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{m\cdot p\cdot e^{-r^{*}}}{q+(m\cdot p)}\right)^{k}=1+q\cdot\left(\frac{m\cdot p\cdot e^{-r^{*}}}{q+(m\cdot p)}\right)+$$

$$\frac{q\cdot(q+1)}{2!}\cdot\left(\frac{m\cdot p\cdot e^{-r^{*}}}{q+(m\cdot p)}\right)^{2}+\ldots=\left[1-\left(\frac{m\cdot p\cdot e^{-r^{*}}}{q+(m\cdot p)}\right)\right]^{-q}=\left(\frac{q+(m\cdot p)-(m\cdot p\cdot e^{-r^{*}})}{q+(m\cdot p)}\right)^{-q}.$$

Variance of the Exponential Dose-Response Function where Dose (K) Has a Delta-Negative Binomial Distribution

$$E[1-e^{-r^*\cdot K}]^2 = \sum_{k=0}^{\infty}[1-e^{-r^*\cdot k}]^2\left[\delta\cdot I_0(k)+(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^k\right]$$

$$=\sum_{k=0}^{\infty}[1-2\cdot e^{-r^*\cdot k}+e^{-2\cdot r^*\cdot k}]\left[\delta\cdot I_0(k)+(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^k\right]$$

$$=\sum_{k=0}^{\infty}[1-2\cdot e^{-r^*\cdot k}+e^{-2\cdot r^*\cdot k}]\left[(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^k\right]$$

$$=(1-\delta)-2\cdot(1-\delta)\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q+\sum_{k=0}^{\infty}[e^{-2\cdot r^*}]^k\left[(1-\delta)\cdot\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\cdot\left(\frac{m\cdot p}{q+(m\cdot p)}\right)^k\right]$$

(from derivation of the expected value of the exponential dose-response function)

$$=(1-\delta)-2\cdot(1-\delta)\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q+(1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\sum_{k=0}^{\infty}\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{(m\cdot p)\cdot e^{-2\cdot r^*}}{q+(m\cdot p)}\right)^k$$

$$=(1-\delta)-2\cdot(1-\delta)\cdot\left(\frac{q}{(q+(m\cdot p))\cdot(1-e^{-r^*})}\right)^q+(1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\cdot\left(1-\frac{(m\cdot p)\cdot e^{-2\cdot r^*}}{q+(m\cdot p)}\right)^{-q}\quad\textit{(see note)}$$

$$=(1-\delta)-2\cdot(1-\delta)\cdot\left(\frac{q}{(q+(m\cdot p))\cdot(1-e^{-r^*})}\right)^q+(1-\delta)\cdot\left(\frac{q}{q+(m\cdot p)}\right)^q\cdot\left(\frac{q+(m\cdot p)-(m\cdot p\cdot e^{-2\cdot r^*})}{q+(m\cdot p)}\right)^{-q}$$

$$=(1-\delta)-2\cdot(1-\delta)\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q+(1-\delta)\cdot\left[\frac{q}{q+(m\cdot p\cdot(1-e^{-2r^*}))}\right]^q$$

$$=(1-\delta)\left[1-2\cdot\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-r^*}))}\right)^q+\left(\frac{q}{q+((m\cdot p)\cdot(1-e^{-2r^*}))}\right)^q\right]$$

From Jolley, 1961, formula 1015:

$$\sum_{k=0}^{\infty}\frac{\Gamma(q+k)}{\Gamma(q)\cdot k!}\cdot\left(\frac{m\cdot p\cdot e^{-2\cdot r^*}}{q+(m\cdot p)}\right)^k=1+q\cdot\left(\frac{m\cdot p\cdot e^{-2\cdot r^*}}{q+(m\cdot p)}\right)+$$

$$\frac{q\cdot(q+1)}{2!}\cdot\left(\frac{m\cdot p\cdot e^{-2\cdot r^*}}{q+(m\cdot p)}\right)^2+\ldots=\left[1-\left(\frac{m\cdot p\cdot e^{-2\cdot r^*}}{q+(m\cdot p)}\right)\right]^{-q}=\left(\frac{q+(m\cdot p)-(m\cdot p\cdot e^{-2\cdot r^*})}{q+(m\cdot p)}\right)^{-q}.$$

APPENDIX J

$$exdr = (1-\delta)\left\{1-\left[\frac{q}{q+(m*p)\cdot(1-e^{-r^{*}})}\right]^{q}\right\}$$

Pr(infection|not infected in the last 30 days) = exdr.
Pr(no infection|not infected in the last 30 days) = (1-exdr).

Pr(infection|infected in the last 30 days) = 1.
Pr(no infection|infected in the last 30 days) = 0.

Let Z represent the number of days a person is infected in a year.

Pr(Z=0) = $(1\text{-exdr})^{365}$.
Pr(Z=1) = Pr(not infected on days 1-364 and infected on day 365) = $(1\text{-exdr})^{364}\cdot\text{exdr}$.
Pr(Z=2) = Pr(not infected on days 1-363, infected on day 364, and remaining infected on day 365) = $(1\text{-exdr})^{363}\cdot\text{exdr}\cdot 1$.

(and so on)

Pr(Z=29) = Pr(not infected on days 1-336, infected on day 337, and remaining infected on days 338-365) = $(1\text{-exdr})^{336}\cdot\text{exdr}\cdot 1^{28}$.
Pr(Z=30) = Pr(infected on day 1, remaining infected until day 30, and not infected on days 31-365; or not infected on day 1, infected on day 2, remaining infected until day 31, and not infected on days 32-365; ... ; or not infected on days 1-335, infected on day 336, and remaining infected until day 365) = $336\cdot(1\text{-exdr})^{335}\cdot\text{exdr}\cdot 1^{29}$.
Pr(Z=31) = Pr(infected on day 1, remaining infected until day 30, not infected on days 31-364, and infected on day 365; or not infected on day 1, infected on day 2, remaining infected until day 31, not infected on days 32-364; and infected on day 365; ... ; or not infected on days 1-334, infected on day 335, remaining infected until day 364, and infected on day 365) = $335\cdot(1\text{-exdr})^{334}\cdot\text{exdr}^{2}\cdot 1^{29}$.
Pr(Z=32) = Pr(infected on day 1, remaining infected until day 30, not infected on days 31-363, infected on day 364, and remaining infected on day 365; or not infected on day 1, infected on day 2, remaining infected until day 31, not infected on days 32-363, infected on day 364, and remaining infected on day 365; ... ; or not infected on days 1-333, infected on day 334, remaining infected until day 363, infected on day 364, and remaining infected on day 365) = $333\cdot(1\text{-exdr})^{333}\cdot\text{exdr}^{2}\cdot 1^{30}$.

In general, the formula for a person being infected z days out of a year is

$$Pr(Z=z)=\frac{(365-z+t)!}{(365-z)!\cdot(t)!}\cdot exdr^{v}\cdot(1-exdr)^{365-z} \qquad z=0,1,\ldots 365$$

where t = largest integer less than or equal to (i/30),
exdr = average probability of infection given in (3), and
v = smallest integer greater than or equal to (i/30).

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=60
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	124	180	235
		25%	165	239	313
		50%	248	359	470
	90th %ile	0%	245	355	464
		25%	326	473	619
		50%	490	709	928
	99th %ile	0%	1,098	1,586	2,070
		25%	1,464	2,115	2,760
		50%	2,196	3,172	4,139

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	13.4×10^{9}
		50%	∞	502,984	84.384
	90th %ile	0%	∞	70.6×10^{6}	697.352
		25%	1.17×10^{9}	212.089	0.700
		50%	32.568	0.098	0.014
	99th %ile	0%	0.007	0.004	0.003
		25%	0.004	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=50
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	129	187	245
		25%	172	250	327
		50%	259	375	491
	90th %ile	0%	312	451	591
		25%	416	602	788
		50%	623	903	1,181
	99th %ile	0%	1,098	1,586	2,070
		25%	1,464	2,115	2,760
		50%	2,196	3,172	4,139

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	1.02×10^9
		50%	24.7×10^{12}	94,107	30.900
	90th %ile	0%	249×10^9	1,754	1.910
		25%	13,775	1.091	0.043
		50%	0.442	0.016	0.006
	99th %ile	0%	0.007	0.004	0.003
		25%	0.004	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=40
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	139	202	264
		25%	186	269	352
		50%	278	404	529
	90th %ile	0%	322	466	610
		25%	429	621	813
		50%	644	932	1,220
	99th %ile	0%	1,098	1,586	2,070
		25%	1,464	2,115	2,760
		50%	2,196	3,172	4,139

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	17.7×10^6
		50%	224×10^9	6,746	6.498
	90th %ile	0%	27.1×10^9	616.588	1.089
		25%	4,392	0.658	0.033
		50%	0.290	0.014	0.005
	99th %ile	0%	0.007	0.004	0.003
		25%	0.004	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=30
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	163	236	309
		25%	217	315	412
		50%	326	472	618
	90th %ile	0%	355	514	673
		25%	474	686	897
		50%	710	1,029	1,346
	99th %ile	0%	1,098	1,586	2,070
		25%	1,464	2,115	2,760
		50%	2,196	3,172	4,139

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	457×10^9
		25%	∞	10.8×10^9	18,872
		50%	38.9×10^6	77.375	0.496
	90th %ile	0%	64.5×10^6	37.963	0.249
		25%	202.563	0.174	0.017
		50%	0.096	0.009	0.004
	99th %ile	0%	0.007	0.004	0.003
		25%	0.004	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=20
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	203	294	385
		25%	271	392	514
		50%	406	588	770
	90th %ile	0%	420	609	796
		25%	561	812	1,062
		50%	841	1,217	1,593
	99th %ile	0%	1,642	2,367	3,082
		25%	2,189	3,155	4,110
		50%	3,284	4,733	6,165

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	12.3×10^{12}	1.02×10^{6}
		25%	∞	130,966	25.222
		50%	5,548	1.018	0.046
	90th %ile	0%	23,236	1.127	0.041
		25%	3.861	0.034	0.008
		50%	0.024	0.005	0.004
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=10
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	255	369	483
		25%	339	492	644
		50%	509	738	966
	90th %ile	0%	548	793	1,037
		25%	731	1,057	1,383
		50%	1,096	1,586	2,074
	99th %ile	0%	3,736	5,344	6,909
		25%	4,982	7,126	9,212
		50%	7,473	10,689	13,818

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	8.69×10^6	212.187
		25%	127×10^6	73.707	0.394
		50%	13.848	0.067	0.011
	90th %ile	0%	8.591	0.043	0.009
		25%	0.086	0.008	0.004
		50%	0.007	0.004	0.003
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=60
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.129	0.187	0.245
		25%	0.172	0.249	0.327
		50%	0.258	0.374	0.490
	90th %ile	0%	0.342	0.495	0.649
		25%	0.455	0.660	0.865
		50%	0.683	0.991	1.298
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=50
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.134	0.195	0.255
		25%	0.179	0.259	0.340
		50%	0.268	0.389	0.510
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=40
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.141	0.205	0.268
		25%	0.188	0.273	0.358
		50%	0.282	0.410	0.537
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=30
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.162	0.235	0.308
		25%	0.216	0.313	0.410
		50%	0.324	0.470	0.615
	90th %ile	0%	0.372	0.540	0.707
		25%	0.496	0.720	0.943
		50%	0.745	1.080	1.415
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=20
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.164	0.238	0.311
		25%	0.218	0.317	0.415
		50%	0.328	0.475	0.622
	90th %ile	0%	0.355	0.514	0.674
		25%	0.473	0.686	0.899
		50%	0.710	1.029	1.348
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=10
WITHOUT ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.164	0.237	0.311
		25%	0.218	0.317	0.415
		50%	0.327	0.475	0.622
	90th %ile	0%	0.356	0.517	0.677
		25%	0.475	0.689	0.902
		50%	0.712	1.033	1.354
	99th %ile	0%	0.710	1.029	1.348
		25%	0.946	1.372	1.797
		50%	1.419	2.058	2.696

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	-∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=60
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	107	155	203
		25%	143	207	271
		50%	214	310	406
	90th %ile	0%	224	325	426
		25%	299	434	568
		50%	449	650	851
	99th %ile	0%	793	1,147	1,499
		25%	1,058	1,529	1,998
		50%	1,587	2,294	2,997

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	448×10^{6}	5,380
	90th %ile	0%	∞	13.6×10^{9}	14,402
		25%	287×10^{9}	3,089	3.052
		50%	280.611	0.256	0.022
	99th %ile	0%	0.043	0.006	0.004
		25%	0.008	0.004	0.003
		50%	0.004	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=50
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	120	173	227
		25%	159	231	303
		50%	239	346	454
	90th %ile	0%	297	430	563
		25%	396	574	751
		50%	594	861	1,127
	99th %ile	0%	793	1,147	1,499
		25%	1,058	1,529	1,998
		50%	1,587	2,294	2,997

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	137×10^{9}
		50%	∞	2.27×10^{6}	209.821
	90th %ile	0%	8.23×10^{12}	9,557	4.796
		25%	86,925	2.496	0.066
		50%	0.873	0.021	0.007
	99th %ile	0%	0.043	0.006	0.004
		25%	0.008	0.004	0.003
		50%	0.004	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=40
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	142	206	270
		25%	190	275	360
		50%	284	412	540
	90th %ile	0%	322	466	610
		25%	429	621	813
		50%	644	932	1,220
	99th %ile	0%	1,101	1,590	2,075
		25%	1,468	2,120	2,767
		50%	2,202	3,180	4,150

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	24.7×10^{12}	6.26×10^{6}
		50%	61.5×10^{9}	3,422	4.369
	90th %ile	0%	27.1×10^{9}	616.588	1.089
		25%	4,392	0.658	0.033
		50%	0.291	0.014	0.005
	99th %ile	0%	0.007	0.004	0.003
		25%	0.004	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=30
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	163	237	310
		25%	218	315	413
		50%	326	473	619
	90th %ile	0%	355	514	673
		25%	474	686	897
		50%	710	1,029	1,346
	99th %ile	0%	1,101	1,590	2,075
		25%	1,468	2,120	2,767
		50%	2,202	3,180	4,150

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	386×10^{9}
		25%	∞	9.26×10^{9}	17,249
		50%	34.7×10^{6}	72.955	0.480
	90th %ile	0%	64.5×10^{6}	37.927	0.249
		25%	202.563	0.174	0.017
		50%	0.096	0.009	0.004
	99th %ile	0%	0.007	0.004	0.003
		25%	0.004	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=20
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	208	301	394
		25%	277	401	525
		50%	415	601	787
	90th %ile	0%	463	669	876
		25%	617	893	1,168
		50%	925	1,339	1,752
	99th %ile	0%	1,642	2,367	3,082
		25%	2,189	3,155	4,110
		50%	3,284	4,733	6,165

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	3.53×10^{12}	368,842
		25%	24.7×10^{12}	53,577	15.173
		50%	2,730	0.730	0.039
	90th %ile	0%	785.963	0.268	0.021
		25%	0.740	0.018	0.006
		50%	0.014	0.005	0.003
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=10
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	290	420	549
		25%	386	560	733
		50%	580	840	1,099
	90th %ile	0%	593	858	1,121
		25%	790	1,144	1,495
		50%	1,186	1,715	2,243
	99th %ile	0%	4,405	6,285	8,107
		25%	5,873	8,380	10,809
		50%	8,810	12,571	16,213

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	24.7×10^{12}	24,844	8.092
		25%	244,127	3.991	0.084
		50%	1.284	0.025	0.007
	90th %ile	0%	1.795	0.024	0.007
		25%	0.042	0.006	0.004
		50%	0.006	0.003	0.003
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=60
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.132	0.192	0.251
		25%	0.176	0.256	0.335
		50%	0.265	0.384	0.503
	90th %ile	0%	0.342	0.495	0.649
		25%	0.455	0.660	0.865
		50%	0.683	0.991	1.298
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=50
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.136	0.197	0.259
		25%	0.182	0.263	0.345
		50%	0.272	0.395	0.517
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=40
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.143	0.207	0.271
		25%	0.190	0.276	0.361
		50%	0.285	0.414	0.542
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=30
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.162	0.235	0.308
		25%	0.216	0.313	0.410
		50%	0.324	0.470	0.615
	90th %ile	0%	0.368	0.533	0.699
		25%	0.491	0.711	0.932
		50%	0.736	1.067	1.398
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs-20
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.162	0.235	0.308
		25%	0.216	0.314	0.411
		50%	0.325	0.471	0.617
	90th %ile	0%	0.349	0.506	0.663
		25%	0.465	0.675	0.884
		50%	0.698	1.012	1.326
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- HAAs=10
WITH ALTERNATE DISINFECTION
JANUARY 2, 1992 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.159	0.230	0.302
		25%	0.212	0.307	0.403
		50%	0.318	0.461	0.604
	90th %ile	0%	0.350	0.508	0.666
		25%	0.467	0.677	0.888
		50%	0.701	1.016	1.331
	99th %ile	0%	0.693	1.006	1.318
		25%	0.925	1.341	1.757
		50%	1.387	2.011	2.635

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=100
WITHOUT ALTERNATE DISINFECTION
DECEMBER 12, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	174	253	331
		25%	233	337	441
		50%	349	506	662
	90th %ile	0%	312	451	591
		25%	416	602	788
		50%	623	903	1,181
	99th %ile	0%	3,723	5,325	6,885
		25%	4,964	7,100	9,179
		50%	7,445	10,650	13,769

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	4.27×10^9
		25%	∞	187×10^6	1,706
		50%	1.67×10^6	16.058	0.206
	90th %ile	0%	252×10^9	1,754	1.912
		25%	13,775	1.091	0.043
		50%	0.442	0.016	0.006
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=75
WITHOUT ALTERNATE DISINFECTION
DECEMBER 12, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	194	280	367
		25%	258	374	490
		50%	387	561	735
	90th %ile	0%	354	512	670
		25%	472	683	894
		50%	707	1,025	1,341
	99th %ile	0%	3,723	5,325	6,885
		25%	4,964	7,100	9,179
		50%	7,445	10,650	13,769

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	10.7×10^6
		25%	∞	1.02×10^6	81.852
		50%	28,134	2.199	0.070
	90th %ile	0%	81.0×10^6	42.112	0.263
		25%	227.327	0.182	0.018
		50%	0.100	0.009	0.004
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=50
WITHOUT ALTERNATE DISINFECTION
DECEMBER 12, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*: 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	240	348	455
		25%	320	463	607
		50%	480	695	910
	90th %ile	0%	419	607	794
		25%	559	809	1,058
		50%	838	1,213	1,587
	99th %ile	0%	3,723	5,325	6,885
		25%	4,964	7,100	9,179
		50%	7,445	10,650	13,769

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*: 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	$220\text{x}10^6$	1,333
		25%	$3.84\text{x}10^9$	376.753	0.955
		50%	51.775	0.120	0.015
	90th %ile	0%	26,626	1.195	0.043
		25%	4.132	0.035	0.008
		50%	0.025	0.005	0.004
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=25
WITHOUT ALTERNATE DISINFECTION
DECEMBER 12, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	257	372	487
		25%	343	496	650
		50%	514	744	974
	90th %ile	0%	548	793	1,037
		25%	731	1,057	1,383
		50%	1,096	1,586	2,074
	99th %ile	0%	3,736	5,344	6,909
		25%	4,982	7,126	9,212
		50%	7,473	10,689	13,818

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	5.45×10^6	163.110
		25%	77.8×10^6	58.329	0.347
		50%	11.454	0.062	0.011
	90th %ile	0%	8.591	0.043	0.009
		25%	0.086	0.008	0.004
		50%	0.007	0.004	0.003
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=100
WITHOUT ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.132	0.191	0.250
		25%	0.175	0.254	0.333
		50%	0.263	0.381	0.500
	90th %ile	0%	0.342	0.495	0.649
		25%	0.455	0.660	0.865
		50%	0.683	0.991	1.298
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=75
WITHOUT ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.144	0.209	0.274
		25%	0.193	0.279	0.366
		50%	0.289	0.419	0.549
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	-∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=50
WITHOUT ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.157	0.227	0.298
		25%	0.209	0.303	0.397
		50%	0.313	0.454	0.595
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=25
WITHOUT ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.165	0.240	0.314
		25%	0.221	0.320	0.419
		50%	0.331	0.480	0.629
	90th %ile	0%	0.358	0.519	0.680
		25%	0.477	0.692	0.906
		50%	0.715	1.037	1.359
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=100
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	171	248	325
		25%	228	331	434
		50%	343	497	650
	90th %ile	0%	294	426	558
		25%	392	568	744
		50%	588	852	1,116
	99th %ile	0%	3,723	5,325	6,885
		25%	4,964	7,100	9,179
		50%	7,445	10,650	13,769

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	13.6×10^{9}
		25%	∞	510×10^{6}	3,084
		50%	3.65×10^{6}	23.657	0.255
	90th %ile	0%	12.3×10^{12}	13,851	5.870
		25%	129,871	2.994	0.072
		50%	1.014	0.022	0.007
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=75
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	196	284	372
		25%	261	379	496
		50%	392	568	744
	90th %ile	0%	354	512	670
		25%	472	683	894
		50%	707	1,025	1,341
	99th %ile	0%	3,723	5,325	6,885
		25%	4,964	7,100	9,179
		50%	7,445	10,650	13,769

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	5.52×10^6
		25%	∞	573,404	58.721
		50%	17,841	1.769	0.062
	90th %ile	0%	81.0×10^6	42.112	0.263
		25%	227.327	0.182	0.018
		50%	0.100	0.009	0.004
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=50
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	239	346	453
		25%	318	461	604
		50%	477	692	906
	90th %ile	0%	419	606	794
		25%	559	809	1,058
		50%	838	1,213	1,587
	99th %ile	0%	3,723	5,325	6,885
		25%	4,964	7,100	9,179
		50%	7,445	10,650	13,769

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	$293\text{x}10^6$	1,574
		25%	$5.21\text{x}10^9$	436.376	1.035
		50%	58.281	0.126	0.016
	90th %ile	0%	26,626	1.195	0.043
		25%	4.132	0.035	0.008
		50%	0.025	0.005	0.004
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=25
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	258	374	489
		25%	344	499	652
		50%	516	747	978
	90th %ile	0%	549	795	1,040
		25%	733	1,060	1,386
		50%	1,099	1,590	2,079
	99th %ile	0%	3,736	5,344	6,909
		25%	4,982	7,126	9,212
		50%	7,473	10,689	13,818

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*		
			0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	$4.45x10^6$	145.502
		25%	$62.8x10^6$	52.690	0.329
		50%	10.547	0.060	0.011
	90th %ile	0%	8.112	0.042	0.009
		25%	0.084	0.008	0.004
		50%	0.007	0.004	0.003
	99th %ile	0%	0.003	0.003	0.003
		25%	0.003	0.003	0.003
		50%	0.003	0.003	0.003

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=100
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.134	0.195	0.255
		25%	0.179	0.259	0.340
		50%	0.268	0.389	0.510
	90th %ile	0%	0.342	0.495	0.649
		25%	0.455	0.660	0.865
		50%	0.683	0.991	1.298
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=75
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*: 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.148	0.214	0.281
		25%	0.197	0.286	0.374
		50%	0.295	0.428	0.561
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1*: 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=50
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.156	0.226	0.296
		25%	0.208	0.301	0.395
		50%	0.311	0.452	0.592
	90th %ile	0%	0.346	0.502	0.657
		25%	0.461	0.669	0.877
		50%	0.692	1.003	1.315
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.

ENHANCED SWTR DATA -- FIRST CUSTOMER -- SURFACE: NO SOFTENING -- TTHMs=25
WITH ALTERNATE DISINFECTION
DECEMBER 16, 1991 VERSION

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Cases per 10,000 per Year	Mean	0%	0.164	0.237	0.311
		25%	0.218	0.316	0.414
		50%	0.327	0.474	0.622
	90th %ile	0%	0.350	0.508	0.666
		25%	0.467	0.677	0.888
		50%	0.701	1.016	1.331
	99th %ile	0%	0.702	1.018	1.334
		25%	0.936	1.358	1.779
		50%	1.404	2.036	2.668

		Secondary Infection Rate	Percentage of Time Log Removal is Reduced by 1* 0%	5%	10%
Average Number of Years for an Outbreak	Mean	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	90th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞
	99th %ile	0%	∞	∞	∞
		25%	∞	∞	∞
		50%	∞	∞	∞

*Statistical method where distribution mean equal to weighted average of resulting finished water levels.