Improved methods for modelling drinking water treatment in quantitative microbial risk assessment; a case study of *Campylobacter* reduction by filtration and ozonation

P. W. M. H. Smeets, Y. J. Dullemont, P. H. A. J. M. Van Gelder, J. C. Van Dijk and G. J. Medema

ABSTRACT

Quantitative microbial risk assessment (QMRA) is increasingly applied to estimate drinking water safety. In QMRA the risk of infection is calculated from pathogen concentrations in drinking water, water consumption and dose response relations. Pathogen concentrations in drinking water are generally low and monitoring provides little information for QMRA. Therefore pathogen concentrations are monitored in the raw water and reduction of pathogens by treatment is modelled stochastically with Monte Carlo simulations. The method was tested in a case study with *Campylobacter* monitoring data of rapid sand filtration and ozonation processes. This study showed that the currently applied method did not predict the monitoring data used for validation. Consequently the risk of infection was over estimated by one order of magnitude. An improved method for model validation was developed. It combines non-parametric bootstrapping with statistical extrapolation to rare events. Evaluation of the treatment model was improved by presenting monitoring data and modelling results in CCDF graphs, which focus on the occurrence of rare events. Apart from calculating the yearly average risk of infection, the model results were presented in FN curves. This allowed for evaluation of both the distribution of risk and the uncertainty associated with the assessment.

Key words | *Campylobacter*, FN-curve, ozone, QMRA, quantitative microbial risk assessment, rapid sand filtration

P. W. M. H. Smeets (corresponding author)
G. J. Medema
Kiwa Water Research, PO BOX 1072,
3430 BB Nieuwegein,
The Netherlands
Tel.: +31 30 6069511
E-mail: *patrick.smeets@kiwa.nl*

Y. J. Dullemont
Waternet, PO Box 8169,
Amsterdam 1005 AD,
The Netherlands

P. H. A. J. M. Van Gelder
J. C. Van Dijk
Delft University of Technology,
PO BOX 5048, 2600 GA Delft,
The Netherlands

INTRODUCTION

Monitoring the absence of indicator organisms in drinking water has been the main approach to safeguard drinking water quality since the beginning of the 20th century (Greenwood & Yule 1917). Drinking water outbreaks have shown that absence of indicator organisms in drinking water does not imply that there is no risk of infection (Hrudey & Hrudey 2004). Since 1980 Quantitative Microbial Risk Assessment (QMRA) has been applied to quantify the microbial safety of drinking water (Haas 1983; Gerba & Haas 1988; Regli *et al.* 1991; Rose *et al.* 1991; Teunis *et al.* 1994; ILSI 1996; Gibson *et al.* 1999; Payment *et al.* 2000). Risk of infection is calculated from the chance of ingesting pathogens (exposure or dose) and the chance of developing an infection from this exposure (dose response relation) (Haas *et al.* 1999). Pathogen concentrations in drinking water are generally below detection limits (Regli *et al.* 1991). QMRA studies have therefore monitored pathogen concentrations in the raw water and modelled removal or inactivation by treatment to

doi: 10.2166/wh.2008.066

estimate concentrations in the drinking water (Teunis *et al.* 1997; Haas & Trussel 1998; Teunis & Havelaar 1999; Westrell *et al.* 2003).

In most of these studies, variability of each element was described by a Probability Density Function (PDF). Treatment was then stochastically modelled by Monte Carlo simulation. Determining the PDF for each element using the available data is a crucial step in such an assessment. PDF parameters have been estimated from pilot study results or literature. However, since raw water concentration and treatment efficacy vary in time and are specific for each drinking water production location, site specific information is preferred (Teunis *et al.* 1997; Nichols 2003; Smeets *et al.* 2007). Monitoring pathogens or indicator organisms in raw water and after treatment steps provides such information. QMRA studies have fitted statistical distributions to such data to determine the PDF. Drinking water risk assessments have mainly used the lognormal, gamma and negative binomial distributions (Teunis *et al.* 1997; Haas *et al.* 1999). Other fields of risk assessment commonly use the Weibull distribution (Van Gelder 1999). The impact of choice of distribution on the result of the risk assessment has not been studied well (Haas 1997). Preliminary studies by the first author indicated that the choice of PDF could dominate the QMRA outcome. Therefore this study focussed on non-parametric techniques for QMRA which do not require a choice of PDF (Van Gelder 1999).

Previous studies (Teunis *et al.* 2004) have shown that extreme events can dominate the average health risk. Historical data on source water concentrations and treatment efficacy can be used to predict normal rare events. These events are caused by the extremes of normal variations in the system such as flow changes, rainfall events, seasonal variations and treatment variations. Observed normal variations are extrapolated to these extreme events by a PDF. Therefore PDF should fit the extremes (tail) of observed variation, in this case monitoring data, since it is used to predict rare events of high concentrations or poor treatment. However, current methods of PDF estimation focus on the distribution type and parameters that best describe the bulk of the data, such as the Kolmogorov-Smirnov test (Haas *et al.* 1999) or likelihood ratio (Teunis *et al.* 1997). This study adopted the use of Complementary Cumulative Distribution Functions (CCDF) graphs (Van Gelder 1999) to visually evaluate the fit of PDF to the tail of the data.

In current QMRA practice the treatment efficacy PDF is validated based on fractions resulting from microbial counts before and after treatment in samples taken on the same date. However, preliminary studies by the first author showed that the predicted concentrations were not in line with the monitored concentrations. Therefore improved methods for model validation were developed in this study.

Much focus in QMRA studies has been on accounting for sampling variability due to (over-) dispersion, variable recovery, pathogen viability and infectivity (Teunis *et al.* 1997; Haas *et al.* 1999; Teunis & Havelaar 1999). The uncertainty that is introduced by Most Probable Number (MPN) type data has not been well studied. Haas *et al.* (1999) treated MPN data similar to count data. Although an 85% correction factor was applied to account for bias in the reported MPN it did not include the uncertainty of the MPN in the risk assessment. Since the case study included MPN type data, a method to include MPN uncertainty was developed.

The outcome of QMRA studies is generally presented as a PDF or histogram of risk of infection (Westrell *et al.* 2003). No distinction between variability and uncertainty was thus made. Other fields of risk assessment such as flooding, traffic or industrial accidents, present societal risk of major accidents in a FN-curve (Van Gelder 1999) plotting the number of casualties (N) *versus* the frequency of occurrence (F). This method seems appropriate for assessment of risk of infection through drinking water, since it is a societal risk. The FN curve allows differentiating between low incidental risk (1 infection per 10,000 people per day) and an outbreak (365 infections per 10,000 people on one day). Although the yearly average risk is identical in both situations, the outbreak is considered less acceptable than the incidental risk. Therefore the FN curve provides better decision support for risk managers and inspectors than a distribution of the yearly average risk.

Methods for large volume sampling, up to 1000 L, have become available in recent years (Hijnen *et al.* 2000; Smeets *et al.* 2007). Since resources are limited, water utilities need to carefully plan their sampling strategy, which includes finding a balance between a limited number of large volume samples and a larger number of regular volume samples.

This study differentiated which concentrations are most relevant for the yearly average risk of infection in order to support such decisions.

The goal of this study was to develop improved methods for modelling drinking water treatment in quantitative microbial risk assessment of drinking water and to apply these methods in a case study. The following methods were adopted from other fields of risk assessment or newly developed:

- Non-parametric bootstrap method for data uncertainty analysis;
- Including MPN uncertainty in the non-parametric bootstrap method;
- Implementation of CCDF graphs for data presentation;
- Verification of validation method (model outcome matches the validation data);
- PDF fitting with focus on tails of data;
- Determination of relative risk related to concentrations;
- Implementation of FN curves for risk presentation.

The paper first describes the methods and the case study and compares different methods of data presentation. Then the non-parametric bootstrap method is applied to determine data uncertainty, including MPN uncertainty. Next the currently applied method to validate pathogen reduction by treatment for Monte Carlo simulation is compared to improved methods. The validated non-parametric treatment model is applied to predict pathogen concentrations after treatment. By comparing the predicted concentrations to the monitored concentrations, the accuracy of the current and improved methods is compared. Next parametric distributions are fitted to the validated model to extrapolate to rare events of high raw water pathogen concentrations or poor reduction by treatment. The predicted concentrations of the parametric treatment model are also compared to the monitored concentrations to verify the accuracy of the model. Risk of infection is then calculated from the concentrations predicted both with the currently applied method and the improved method of treatment model validation. Risk of infection is determined for each concentration to assess the relative impact on the risk, which provides guidance for monitoring. Finally the risk assessed with current and improved methods is compared in a FN curve.

METHODS

Case description

Campylobacter monitoring data collected at the water treatment plant (WTP) Leiduin of Waternet (water cycle company for Amsterdam and surrounding areas) was used for the case study. The source for the WTP is water from the river Rhine which is pre-treated and infiltrated in the dunes. The water is abstracted in open canals and collected in an open reservoir before post-treatment. The pre-treatment and the soil passage remove most pathogens from the Rhine water; however the water is re-contaminated by birds and wildlife in the open canals and reservoir directly before the treatment plant intake. The water in the reservoir was referred to as raw water in the case study. The reservoir is situated in a protected dune area, therefore contamination of the reservoir through waste water and agricultural run-off is unlikely. Water fowl like ducks, geese, gulls and swans in the reservoir and the abstraction canals are the most likely sources of *Campylobacter*. Wildlife like deer, rabbits and rodents and possibly some pets (cats, dogs) in the dunes can also contribute to the contamination with *Campylobacter*. Contamination can take place either by entering the water (rats, dogs) or by shedding faeces on the shore, which is then washed into the reservoir by rain during run-off. The reservoir is refreshed daily. Since the contamination takes place only in a small proportion of the water (the water surface where the ducks are swimming and the shores) and the reservoir is not mixed, the water quality at the intake sampling point is likely to vary significantly. The raw water is treated by rapid sand filtration, ozonation, softening, biological activated carbon filtration and slow sand filtration. Rapid sand filtration, ozonation and slow sand filtration are considered the main microbial barriers at the WTP (Figure 1). The risk of infection was calculated for consumption of ozonated water. Since some *Campylobacter* were detected

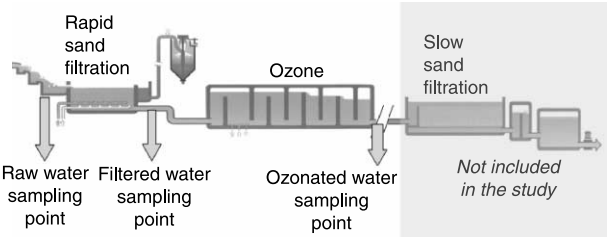

Figure 1 | Microbial barriers and location of sampling points at WTP Leiduin.

in ozonated water this dataset was more appropriate to demonstrate the improved methods than the drinking water dataset in which no *Campylobacter* was detected.

Microbial analysis

Campylobacter samples were analyzed by direct filtration and direct inoculation of the filter in tubes with Preston *Campylobacter* selective enrichment broth. Positive results were confirmed by microscopic examination of a hanging drop for the presence of *Campylobacter*. *Campylobacter* was quantified by the most probable number method (MPN) in three parallel tubes for three filtered sample volumes using decimal dilutions. The collected monitoring data consisted of the MPN-arrays for each sample (e.g. 3-2-1 indicated three positive tubes in the largest volume, two positive tubes in the middle volume and one positive tube in the smallest volume). Reported MPNs were taken from MPN tables by De Man (1975).

Non-parametric MPN bootstrapping

The bootstrap method is a fairly easy tool to numerically calculate the uncertainty of a dataset of measurements, by repeatedly drawing results randomly from the dataset. The confidence interval for the monitored *Campylobacter* concentrations was determined by adapting a standard non-parametric bootstrapping procedure (Van Gelder 1999) to include MPN method uncertainty (De Man 1975). A result was randomly drawn from the m monitoring results (with replacement, Equation (1) and for this result a random concentration was drawn according to the MPN likelihood distribution for the result (Equation (2)). Thus a bootstrap sample of *Campylobacter* concentrations in the monitored water was produced.

$$X^*_{ij} = X_{[m\cdot p]} \quad i = 1, \ldots, n \quad j = 1, \ldots, k \tag{1}$$

$$q(C^*_{ij}|X^*_{ij}) = (1 - e^{-v_1 C_{ij}})^{P_1}(1 - e^{-v_1 C_{ij}})^{3-P_1}(1 - e^{-v_2 C_{ij}})^{P_2} \ldots (1 - e^{-v_2 C_{ij}})^{3-P_2}(1 - e^{-v_3 C_{ij}})^{P_3}(1 - e^{-v_3 C_{ij}})^{3-P_3} \tag{2}$$

Where X is the dataset of monitored MPN results, X^* is the bootstrap dataset of MPN results, i indicates the ith MPN result, j indicates the jth bootstrap sample, n is the number of draws in a bootstrap sample, k is the total number of bootstrap samples in a bootstrap dataset, p and q are uniform random variables, *[m.p]* is the integer ceil function (round up of *m.p*), C^* is the bootstrap sample of *Campylobacter* concentrations (organisms/L), q is the likelihood of concentration C^* given result X^*, v_1, v_2, v_3 are the three volumes (or dilutions) used in the MPN method and P_1, P_2, P_3 are the number of positives at the respective volumes. Equation (2) was solved numerically to determine C^* at a given q and X^*. By producing k bootstrap samples of size n with $n = m$ the C^* resembled the likelihood of *Campylobacter* concentrations given the presence/absence results. From this the 95% confidence interval (CI) of the concentration was determined for each proportion of the water. Stable results were achieved with acceptable calculation times for $k = 10{,}000$.

In some ozonated water samples no *Campylobacter* were detected (0-0-0 result). Consequently the MPN likelihood q in Equation (2) approaches 1 as C approaches 0, so no lower limit of the likely concentration can be determined. As a practical approach, the non detects were adapted before bootstrapping by doubling the sample volume and assuming one positive in the largest MPN volume (1-0-0 result). This is similar to setting non-detect samples of count data to 'half the detection limit', which is a conservative approach in risk assessment. Similarly, raw water samples that were all positive (3-3–3 result) were adapted to one negative in the smallest MPN volume (3–3–2 result) with half the sample volume to provide an upper limit of likely concentration. Although this is a simplified approach for these 'larger than' values, it proved to be efficient to demonstrate the methods in this study. Preferably these issues are prevented during monitoring by using sufficient sample dilutions. The bootstrap samples of raw, filtered and ozonated water were used for the assessment of pathogen reduction by treatment, the assessment of the raw water PDF and model verification.

Non-parametric validation of treatment efficacy

Treatment efficacy π is the fraction of organisms which pass a treatment step. The observed treatment efficacy was calculated from the bootstrap datasets of monitoring data as:

$$\pi^*_{ij} = \frac{C^*_{out\,[n\cdot p1][k\cdot p2]}}{C^*_{in\,[n\cdot p1][k\cdot p3]}} \quad i = 1, \ldots, n \quad j = 1, \ldots, k \tag{3}$$

where *p1*, *p2* and *p3* are uniform random variables and *[n.p1]* is the integer ceil function. Several methods can be used

to select values from the bootstrap samples C^*_{in} and C^*_{out} that are 'paired' in Equation (3). The effects of using either the 'random', 'date' or 'rank' method were studied. The bootstrap samples C^*_{in} and C^*_{out} require different preparations for these methods.

The random method assumes no correlation by date or rank. The bootstrap samples C^*_{in} and C^*_{out} did not undergo any adaptation, so samples before and after treatment were paired randomly (since samples X^* were selected randomly in Equation (1)).

Pairing by date has been widely applied in QMRA (Teunis *et al.* 1997; Teunis & Havelaar 1999; Teunis *et al.* 1999) and can be considered the current 'state of the art'. Influent and effluent samples taken on the same day are compared and π is calculated for each pair. This assumes that samples before and after treatment are correlated in time. To enable pairing by date, the monitoring datasets X_{in} and X_{out} were prepared so that they only included results taken on the same day in date order. Equation (1) was adapted so that samples were drawn in order and without replacement (*[m.p]* was replaced by i so all bootstrap samples included every result once). Effectively the date-bootstrap procedure produced a bootstrap dataset which only included MPN uncertainty.

Pairing by rank has only been reported once (Teunis *et al.* 1999) and was referred to as 'unpaired counts', but its application was not explored further. Pairing by rank assumes complete correlation between the influent and effluent concentrations (lowest influent concentrations correlate to lowest effluent concentrations etc.). To enable pairing by rank, the bootstrap samples C^*_{in} and C^*_{out} were sorted by concentration before determining π.

Using Equations (1), (2) and (3), π^*_{filt} was determined from C^*_{raw} and C^*_{filt}, and π^*_{O3} was determined from C^*_{filt} and C^*_{O3}. Thus π^* resembled the likelihood of actual *Campylobacter* reduction by removal and inactivation respectively. From this the 95% confidence interval (CI) of the reduction was determined for each proportion of the water for presentation in graphs. The study used the total bootstraps in calculations, not the 95% CI.

Parametric extrapolation of bootstrap samples

Parametric distributions were fitted to the k bootstrap samples of n raw water C^*_{raw}, removal π^*_{filt} and inactivation π^*_{O3} values respectively using the fit functions in Matlab® for several distribution types. This resulted in k parameter pairs for each distribution type. Gamma, lognormal and Weibull distributions were fitted to C^*_{in}, π^*_{filt} and π^*_{O3}. The beta distribution was only fitted to π^*_{filt} and π^*_{O3}.

$$G_j = PDFfit(C^*_j) \text{ respectively } H_j = PDFfit(\pi^*_j)\ j = 1, \ldots, k \quad (4)$$

Where G_j is the parameter pair of the PDF fitted to the concentration bootstrap sample C^*_j, H_j is the parameter pair of the PDF fitted to the reduction bootstrap sample π^*_j, and *PDFfit* is the fit function in Matlab® for the chosen PDF type.

Non parametric treatment model

Monte Carlo simulation was used to model reduction of pathogens by treatment. By using the bootstrap samples of C^* and π^* in Equation (5) a non-parametric model of *Campylobacter* reduction by treatment was achieved. This model was used to verify which of the methods (random, date or rank) provided the best validation for Monte Carlo simulations. The number of draws in one simulation n was set to m (the number of monitoring samples) to verify whether the model predicted the distribution of concentrations after treatment correctly. The validation was considered to be correct when the predicted concentrations after treatment overlapped the monitored concentrations.

$$C^*_{\text{out\_}ij} = C^*_{in[n\cdot p1][k\cdot p2]}\,\pi^*_{[n\cdot p3][k\cdot p4]} \quad i = 1, \ldots, n \quad j = 1, \ldots, k \quad (5)$$

Parametric treatment model

To predict the likelihood of rare events of high concentrations, Monte Carlo simulation with the parametric PDFs (G for the raw water concentration and H for the reduction) was applied as:

$$C^{\#}_{\text{out\_}ij} = PDFrnd(G_{[k.p1]})PDFrnd(H_{[k.p2]}) \quad i = 1, \ldots, n \quad j = 1, \ldots, k \quad (6)$$

Where $C^{\#}_{out\_ij}$ is the predicted concentration after the treatment step, *PDFrnd* is the random draw of realisations from a given PDF function in Matlab®, $G_{[k.p1]}$ and $H_{[k.p2]}$ are a random PDF parameter pair of the raw water and the

reduction respectively. The number of simulations n was chosen with respect to the proportion of time that was of interest (i.e. $n = 10{,}000$ was applied in this study to predict events which can occur up to 0.01% of the time).

Risk calculation

Exposure was calculated from the *Campylobacter* concentration in the drinking water and consumption of unboiled drinking water. For QMRA purposes the consumption can also be modelled as a PDF. However for this study only the average consumption was used since the goal was to show the impact of treatment modelling (using a PDF for consumption would distort these effects). Daily exposure (dose) μ_d (*Campylobacter*/d) was calculated by multiplying the estimated concentration with the average Dutch consumption of 0.177 litre of unboiled drinking water per day (Mons *et al.* 2007). The daily risk of infection P_{inf_d} (infection per person per day) was calculated from exposure using a Beta-Poisson dose-response model for *Campylobacter* with $\alpha = 0.145$ and $\beta = 7.59$ (Medema *et al.* 1996).

$$P_{\inf\_d} \approx 1 - \left(1 + \frac{\mu_d}{\beta}\right)^{-\alpha} \quad (\beta \geq 1 \text{ and } \alpha \leq \beta) \tag{7}$$

Since the concentration in the drinking water varies in time, the exposure also varies in time. In theory a frequently occurring low concentration could pose the same average yearly health risk as a rarely occurring high concentration. To assess the relative impact of occurring concentrations, the yearly risk of infection from exceeding a given concentration C_i (organisms/L) for a proportion of the year F_i (dimensionless) was calculated with Equation 8. Yearly risk of one or more infections P_{inf_y} (infection per person per year) was calculated with $F_i = 1$.

$$P_{\inf\_y\_i} = 1 - (1 - P_{\inf\_d\_i})^{365F_i} \tag{8}$$

RESULTS

Microbial monitoring

Samples were taken at the raw water sampling point, mixed filter effluent and ozonation effluent. *Campylobacter* was analysed monthly in 2003 and 2005. In the winter period (December to February) of 2003, 2004 and 2005 *Campylobacter* analysis was performed weekly. Table 1 provides an overview of the collected data. Figure 2 shows the sample results on a time scale including the uncertainty due to the MPN method.

Methods to present distribution of concentrations

The variation of *Campylobacter* concentration in time needs to be taken into account for QMRA. Currently monitoring data is presented in QMRA studies as histograms to fit a PDF or cumulative histograms to fit a CDF on a semi-log scale. In this study the data was presented as Complementary Cumulative 'Histogram' to fit a Complementary Cumulative Distribution Function (CCDF) on a double log scale. This form of presenting data is generally applied in other fields of risk assessment and is well suited for extrapolation to rare events. Since the proportion of samples (similar to frequency) is plotted on log scale, and 'rare' events occur a small proportion of the time, this part of the data is 'magnified'. Figures 3(a), 3(b) and 3(c) show the raw water monitoring data as PDF, CDF and CCDF respectively.

Figure 4 shows the CCDF of the monitored *Campylobacter* MPN numbers in raw, filtered and ozonated water. It also shows the median and 95% confidence interval, as determined with non-parametric MPN bootstrap. Since 30 to 40 *Campylobacter* samples were taken at each sampling point, each sample represents a proportion of 2.5–3.3% of the produced water.

Table 1 | Overview of *Campylobacter* monitoring data in Most Probable Number/L (MPN/L)

	#	Mean MPN/L	Median MPN/L	Min MPN/L	Max MPN/L	St. Dev.	Skew	Kurtosis
Raw	41	197	110	0.30	1,100	224	1.90	7.53
Filtered	32	11.6	4.0	0.40	110	20.6	3.72	17.65
Ozonated	31	0.04	<0.03	<0.03	0.40	0.10	3.26	12.26

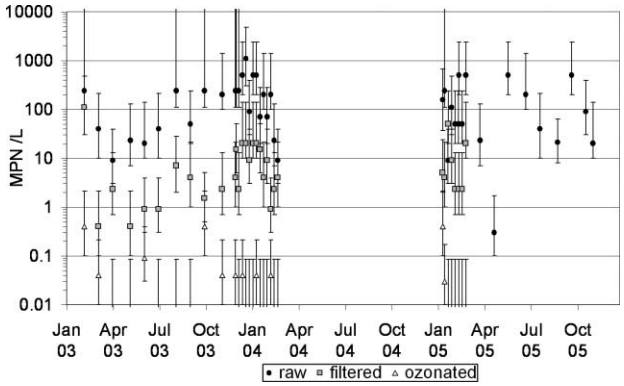

Figure 2 | *Campylobacter* monitoring results in raw water, filtered water and ozonated water. Error bars indicate the 95% confidence interval of the MPN for *Campylobacter*.

Non-parametric treatment model

The non-parametric stochastic model of treatment efficacy was validated with the *Campylobacter* monitoring data (Table 1 and Figure 4) by non-parametric validation of treatment efficacy. Figures 5(a) and 5(b) show the estimated *Campylobacter* removal by filtration π^{*}_{filt} using the random, date and rank method. The removal found with the random method showed the highest variability of treatment efficacy. The date method resulted in similar removal, so pairing samples by date had little impact on the estimation of removal. Both the random and the date method results allowed for 'negative removal' to occur ($\pi_{filt} > 1$). This would imply that pathogens were sometimes "produced" by the filter, which is unlikely. The rank method resulted in approximately 1 log removal and little variation. The rank method did not allow for negative removal.

Figures 6(a) and 6(b) show the estimated inactivation of *Campylobacter* by ozonation. The random and the date method resulted in a very similar estimate of inactivation by ozonation. Both showed high variability of inactivation and possible occurrence of negative inactivation. The rank method resulted in a more stable inactivation of approximately 2 log.

The estimated removal and inactivation in the previous section provided the non-parametric validation of the stochastic treatment model. The impact of the method of validation on the predicted concentrations after a treatment step was determined for the currently applied date method and the new rank method. The random method was not included in the rest of the study since the results were very similar to the date method. The concentrations after filtration were calculated from the raw water bootstrap samples (Figure 4) and the validated removal using the date method or the rank method (Figure 5(b)). The calculated concentrations were compared to the bootstrap of filtered water monitoring results C^{*}_{filt} in Figure 7(a) (date method) and 7(b) (rank method). The concentrations after ozonation were calculated from the filtered water bootstrap samples (Figure 4) and the validated inactivation using the date method or the rank method (Figure 6(b)). The calculated concentrations were compared to the bootstrap of ozonated water monitoring results C^{*}_{O3} in Figure 8(a) (date method) and 8(b) (rank method).

Figures 7(a) and 8(a) show that the date method resulted in substantial over-estimation of *Campylobacter* concentrations both after filtration and ozonation. The rank method

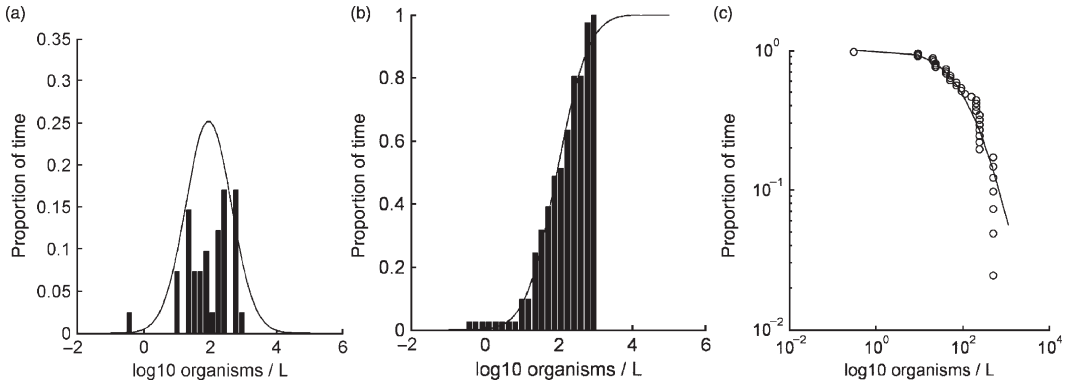

Figure 3 | Distribution of raw water *Campylobacter* monitoring data (bars or markers) and fitted lognormal distribution (lines) as PDF (3(a)), CDF (3(b)) and CCDF (3(c)).

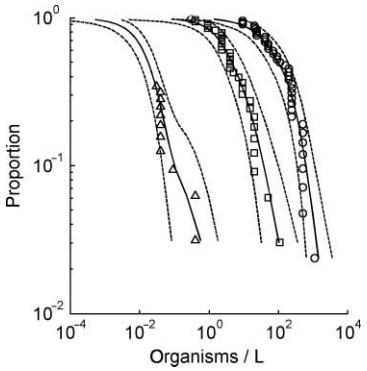

Figure 4 | CCDF of monitored *Campylobacter* MPN concentrations (markers) and the medians (lines) and 95% CI (dashed lines) of the non-parametric bootstraps for raw water (○), filtered water (□) and ozonated water (△).

provided an appropriate estimate of π_{filt} and π_{O3} for Monte Carlo simulation since the monitored concentrations in Figures 7(b) and 8(b) are in line with the predicted concentrations. The rank method was used in the rest of the study since it provided the best validation of the treatment model. The date method was included to demonstrate the error caused by this currently used method.

Parametric treatment model

The non-parametric model cannot predict rare events of high *Campylobacter* concentrations or poor treatment removal due to the limited number of samples. The non-parametric model validations C^*_{raw}, π^*_{filt} and π^*_{O3} were therefore extrapolated with parametric distributions. Figures 9(a) and 9(b) show that both the Weibull and gamma distribution under-estimated rare high *Campylobacter* concentrations in raw water. This would result in underestimating the risk of infection from rare events. The lognormal distribution in Figure 9(c) matched the shape of C^*_{raw} and was therefore chosen to extrapolate the raw water *Campylobacter* concentrations in this study.

The obtained bootstraps of reduction by treatment π^*_{filt} and π^*_{O3} were extrapolated to rare events of poor reduction (high values of π) with the Weibull, beta, gamma and lognormal distribution. Figure 10(a) shows the fit of the gamma distribution to π^*_{filt}. Weibull and beta distributions provided a practically identical graph and are therefore not shown. Figure 10(b) shows the fit of the lognormal distribution to π^*_{filt}. Although all distributions provided a reasonable fit for most of the data, only the lognormal distribution provided a reasonable fit to the high π^*_{filt} values (poor removal). Therefore the lognormal distribution was used in further analysis in this study.

The gamma and lognormal distributions were fitted to the bootstrap of inactivation by ozone π^*_{O3}. Again the lognormal distribution provided the best fit to rare events of poor inactivation. The lognormal distribution was therefore used in the rest of the study.

Parametric model of total chain

Monte Carlo simulation of the treatment from raw to ozonated water was performed to estimate the occurrence of *Campylobacter* in ozonated water. The parametric model was used to include normal rare events. The concentration in raw water, removal by filtration and inactivation by ozonation were modelled with lognormal distributions. Figures 11(a) and 11(b) show the resulting median and 95% CI of predicted *Campylobacter* concentrations at each step compared to the monitored concentrations and their 95% CI. The date method shown in Figure 11(a) predicts

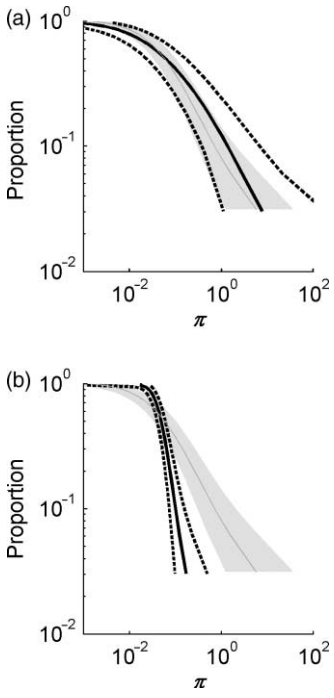

Figure 5 | Non-parametric validation of *Campylobacter* removal by filtration π^*_{filt} with the date method (grey area) compared to random (5(a)) and rank (5(b)) method, median (line) and 95% CI (dashed).

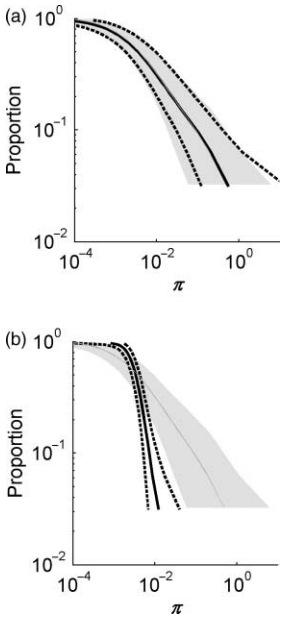

Figure 6 | Non-parametric validation of *Campylobacter* inactivation by ozonation π^{*}_{O3} with the date method (grey area) compared to random (6(a)) and rank (6(b)) method, median (line) and 95% CI (dashed).

concentrations after filtration and ozonation which are very high compared to the monitoring results. This indicates that currently applied QMRA methods based on pairing monitoring data by date can significantly over-estimate the concentration of pathogens after treatment. The new method of pairing by rank resulted in a stochastic model of treatment which predicts concentrations in line with monitoring results (Figure 11(b)). Since the same data was used for validation and verification, this study only demonstrated that the rank method results in an accurate model, whereas the date method overestimated concentrations. The predictive accuracy of the rank method will be assessed in a subsequent study by using separate datasets for validation and verification.

Modelled risk of infection

The risk of infection from consuming ozonated water was calculated based on the modelled concentration in ozonated water. The choice of method to determine reduction by treatment had a significant impact on the assessed risk. The individual health risk is represented by the average yearly risk of infection. The date method predicted a 70% (33%–96%) average yearly risk of infection, whereas the rank method predicted 8.3% (3.8%–18%). So the conventional date method predicted a ten times higher average yearly risk of infection than the new rank method. The Dutch drinking water guidelines (Anonymous 2001) require a maximum individual risk of 10^{-4} yearly average risk of infection, which corresponds to $2.75^{*}10^{-7}$ daily risk of infection. Approximately 3 log reduction was needed in order to achieve this level of safety in the drinking water. The slow sand filtration at WTP Leiduin further treated the ozonated water to achieve this reduction.

Figure 12(a) shows that according to the date method the risk was dominated by concentrations of approx. 28 *Campylobacter*/L (black line) occurring in 1% of the water (grey line), which corresponds to an average yearly risk of 25% (black line). This concentration is 70 times higher than the maximum monitored concentration of 0.4 *Campylobacter* /L in Table 1 observed in 3% of the samples. Figure 11(a) however shows that the concentration after ozonation

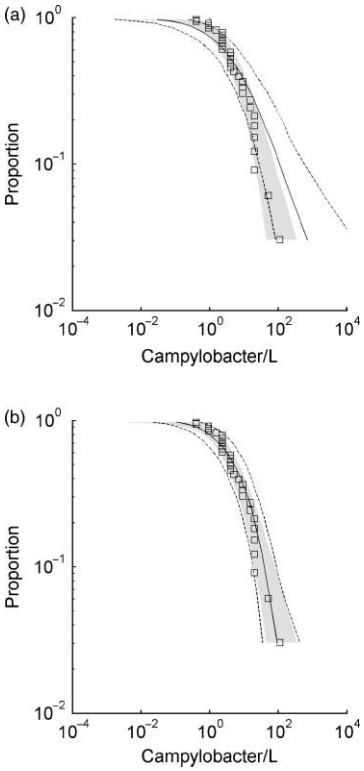

Figure 7 | *Campylobacter* concentration in filtered water calculated with the non-parametric model validated by conventional date method (7(a)) and new rank method (7(b)), median calculated concentration (line) and 95% CI (dashed), compared to monitored concentrations (markers) and 95% CI of C^{*}_{filt} (grey area).

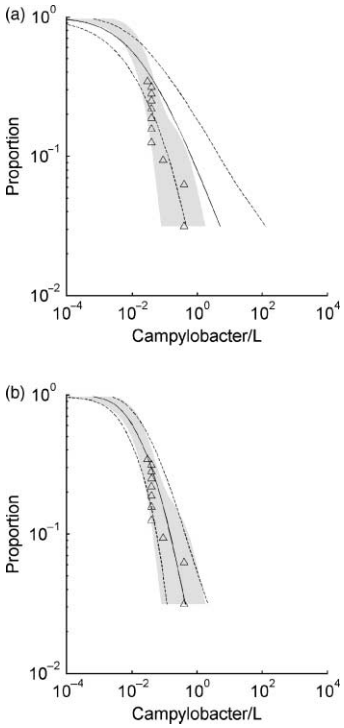

Figure 8 | *Campylobacter* concentration in ozonated water calculated with the non-parametric model validated by conventional date method (8(a)) and new rank method (8(b)), median calculated concentration (line) and 95% CI (dashed), compared to monitored concentrations (markers) and 95% CI of C^{*}_{O3} (grey area).

predicted with the date method is not in line with the monitoring data therefore this high estimate of risk seems unlikely.

Figure 12(b) shows that according to the rank method the average yearly risk was dominated by concentrations of approx. 0.14 *Campylobacter*/L (black line) occurring in 10% of the water (grey line), which corresponds to an average yearly risk of 1.7% (black line). This concentration was exceeded in 10% of the monitoring samples; therefore the estimate of the frequency was regarded accurate. The extrapolation through modelling predicted that higher concentrations did not have a significant impact on the average yearly health risk.

Since the choice of treatment model validation method appears to have a significant effect on the assessed risk, the model results need to be compared to the original monitoring data. The modelling also provides guidance for future monitoring. Frequent sampling of 10 L volumes will verify or improve the estimate of the concentrations which dominate the risk of infection. Lower concentrations which required larger volumes have little effect on the risk estimate. Smaller sample volumes would result in negative samples only, thus providing no additional information for the QMRA.

The FN curve of daily risk of infection is shown in Figure 13(a) and 13(b). Filtration had a limited effect on the daily risk whereas ozonation had a major impact. The date method (Figure 13(a)) predicted more frequent occurrence of high risk than the rank method (13(b)). The rank method provided the best validation of the model, therefore only the FN curve for the rank method (Figure 13(b)) is discussed here. The FN curve shows both the variation of risk and the uncertainty of the assessed risk thus supporting decisions by risk managers and inspectors. The societal risk can be

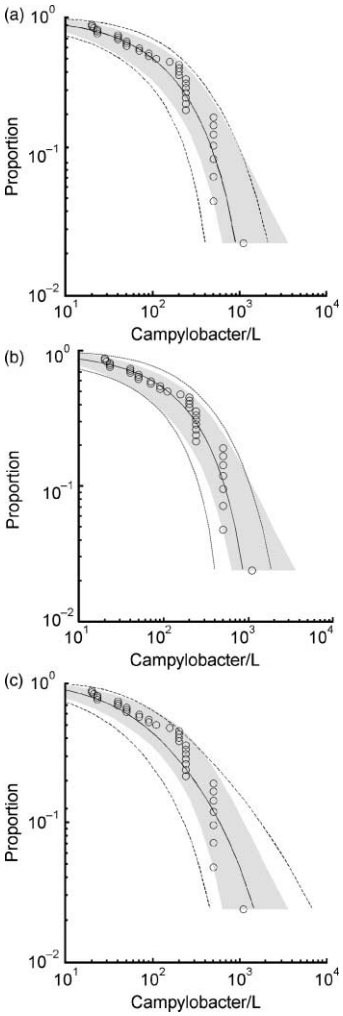

Figure 9 | Median (line) and 95% CI (dashed) of Weibull (9(a)), gamma (9(b)) and lognormal (9(c)) distributions fitted to the non parametric bootstrap (95% CI in grey) of *Campylobacter* concentrations in raw water. Markers indicate the monitored concentrations in MPN/L.

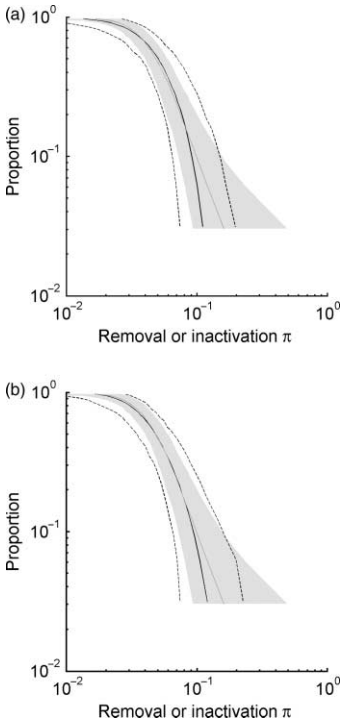

Figure 10 | Median (line) and 95% CI (dashed) of gamma (10(a)) and lognormal (10(b)) distributions fitted to the non parametric bootstrap of *Campylobacter* removal by filtration π^{*}_{filt} (grey area indicates 95% CI, grey line indicates median). Weibull and beta distributions (not shown) provided a graph identical to the gamma distribution (a).

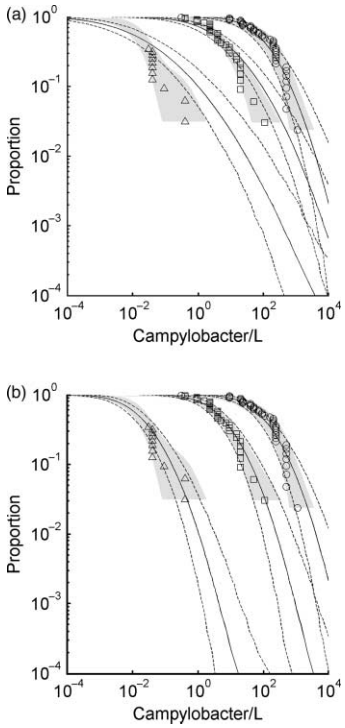

Figure 11 | Monte Carlo simulation of *Campylobacter* concentrations at different stages of treatment validated with the date (11(a)) and rank (11(b)) method, median (lines) and 95% CI (dashed), compared to monitored concentrations (markers) and 95% CI of monitoring (grey area) at several stages in treatment; raw water (○), filtered water (□) and ozonated water (△).

evaluated with the FN curve by evaluating the likelihood of simultaneous infection of a large number of people, referred to as an outbreak. An outbreak is represented in the FN curve by a high daily risk of infection. The FN curve in Figure 13(b) shows that the risk of infection from drinking ozonated water exceeds 0.7% one day per year (proportion of 0.0027). In a city of 1 million people 7,000 people would gain an infection of which some would develop illness. The upper 97.5 confidence limit of this estimate is 2% risk of infection one day per year, resulting in 20,000 infections. Since outbreaks may be detected when over 1% of the population becomes ill (Regli *et al.* 1991), the risk assessment indicates that an outbreak might be detected yearly for this case study. A detected outbreak would result in a much greater effect on society than the incidental infections due to the yearly average risk. Current legislation does not set requirements for the acceptable frequency and magnitude of such an outbreak. The numbers in this example are hypothetical since the ozonated water passes slow sand filtration before distribution which reduces the risk.

DISCUSSION

Monte Carlo simulation of treatment is common practice in current QMRA studies. Since removal by treatment cannot be measured directly, it is calculated from concentrations before and after treatment measured on the same day. This approach assumes a correlation in time between these individual samples. However, it is known that such correlation is disturbed by several causes, even when the residence time in the treatment process is accounted for. Firstly, sampling variation due to (over-)dispersion of organisms in the water needs to be accounted for. Gale *et al.* (1997) showed that treatment enhances clustering of micro-organisms, thus impacting the dispersion. Secondly, the residence time of particles in some processes (e.g. filtration) can be very different from the water residence time (Yao *et al.* 1971). In addition, treatment processes vary in time (filtration cycles) and in space (inhomogeneous mixing of disinfectants). Finally microbial methods can have a large impact due to the quantification uncertainty (MPN, presence/absence) or

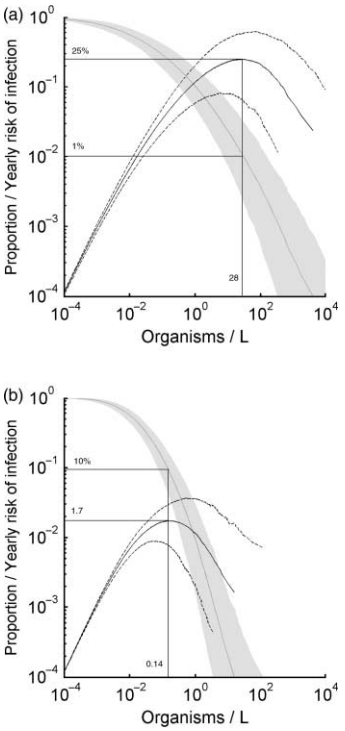

Figure 12 | Median (grey line) and 95% CI (grey area) of modelled *Campylobacter* concentration in ozonated water, and median (line) and 95% CI (dashed) of yearly risk of infection related to the proportion of each concentration using the date (12(a)) and the rank (12(b)) method.

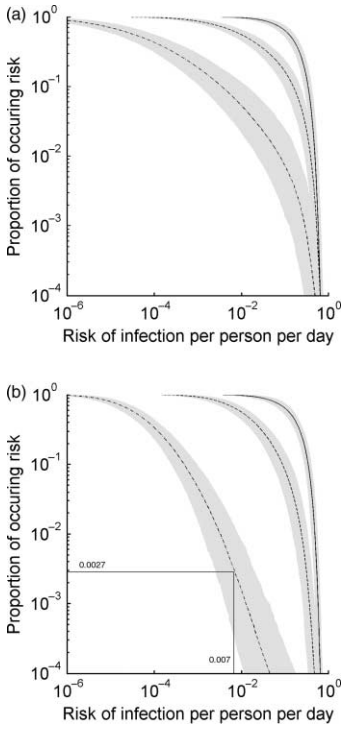

Figure 13 | FN curve of median (lines) and 95% CI (grey area) of daily risk of infection from drinking raw water (line), filtered water (dashed) or ozonated water (dash-dot) using the date (13(a)) and the rank (13(b)) method.

recovery. Several methods have been published to account for these disturbances such as statistical correction for recovery (Teunis & Havelaar 1999) or the use of copula's or correlations (Bukowski *et al.* 1995; Haas 1999). In this case study, pairing by date resulted in the same assessed removal as random pairing, indicating that there is little correlation in time between influent and effluent data. Predicted concentrations after filtration and ozonation deviated strongly from the monitored concentrations. Since the treatment model was not able to predict its validation data, it can be concluded that it was not validated correctly. The newly developed rank method proved to be very effective for model validation. This method implies that samples taken years apart may be paired, which contradicts to the intuitive expectation that only samples taken within a short time frame may be correlated. However, one must consider that the goal of the Monte Carlo simulation is to model the transition from the raw water distribution to the treated water distribution, not to predict the chance of an individual micro-organism passing treatment.

The presented results were obtained for one case study, the applicability to other situations needs to be studied further. Since correlation in time may be relevant for other treatment systems, the choice of date or rank method must always be made with care. This study provided two methods for this. Firstly the random method provides a benchmark for data with no correlation. When the date method results in a significant deviance from the random method, this indicates that correlation in time has a significant effect. Secondly, concentrations after treatment predicted by non-parametric modelling should be in line with the validation data, taking into account the uncertainty of a limited dataset and method uncertainties.

Reported removal by treatment in literature is also applied in QMRA. Generally reported removal ranges over several log units, so the choice of removal in a QMRA study will significantly impact the assessed risk. One needs to consider that the date method was generally used to determine these literature values of removal. The results from this study lead to a new consideration of the reported data, since the rank method could lead to significant reduction of the range of reported removal.

The study adapted methods which are generally applied in other fields of risk assessment, such as flooding, traffic or industrial accidents (Van Gelder 1999), for application in drinking water for QMRA. This includes the use of CCDF, non-parametric bootstrapping and the FN curve. The main difference is that for many other fields of risk the extremes (water levels, fatal accident, earth quakes or process temperatures) can be monitored directly, leading to other extrapolation techniques such as peak over threshold (POT). It is unlikely that microbial monitoring catches the actual peak contaminations or moments of poor treatment. Microbial monitoring results must therefore be considered as random samples of the variation, to be extrapolated with statistical distributions. Adapting monitoring strategies to capture the real peaks may provide a significant improvement of the assessed risk. It must also be considered that the techniques presented in this study only predict the events due to (combinations of) 'normal' variations. Assessment of other 'man-made' events, such as operational errors or intentional contamination, need to be addressed with other methods, such as water safety plans (WHO 2004).

Currently the individual risk, expressed as average yearly risk of infection or DALY, is the main parameter for risk evaluation (WHO 2004). The prevention of outbreaks however is one of the main concerns of water utilities and health authorities. The FN curve allows for evaluation of both the individual risk and the societal risk of 'outbreaks'. Further more it provides the uncertainty involved for both these aspects. The FN curve thus provides a basis for a new approach to risk evaluation and legislation.

Microbial monitoring remains important to verify the achieved level of safety. This study provided a method to determine the concentrations which are most relevant for the yearly average risk of infection. This can support monitoring programs in order to efficiently direct resources e.g. by taking frequent small volume samples, rather than a few large volume samples. Since the presented methods assume 'random' samples, a large volume sample cannot be considered as a large number of small samples. After all, it cannot be assumed that the distribution of concentrations in the large volume is identical to the distribution in the yearly produced water. Still large volume samples may be necessary to get a relevant number of positive samples per year. This means that sampling strategy may need to be adapted based on monitoring results: first find positives, and then determine concentrations most relevant for risk.

CONCLUSIONS

The currently applied method to model drinking water treatment in QMRA was compared to an improved method. This study showed that the currently applied method did not predict the monitoring data used for validation in a case study with *Campylobacter* monitoring data of filtration and ozonation processes. Consequently the risk of infection was over estimated by one order of magnitude in this case study. The improved method accurately predicted the validation data. In this case the rank method proved to be the best validation method, however this may not be the case for all systems. The study also introduced other techniques to QMRA which improve calculation, presentation and evaluation of data and risk. Since CCDF graphs focus on rare events, visual evaluation of modelled extrapolation is improved. The use of non-parametric methods prevents the impact PDF choice in an early stage of QMRA. Calculating the risk per concentration provides guidance for monitoring and the FN curve allows improved risk evaluation by distinguishing between individual and societal risk. Together these methods provide an improved protocol for modelling drinking water treatment in QMRA.

ACKNOWLEDGEMENTS

The research was funded by the MicroRisk project and the joint research program of the Dutch water companies. The MicroRisk (*www.microrisk.com*) project is co-funded by the European Commission under contract number EVK1-CT-2002-00123. The authors wish to thank Dr. S.R. Petterson from the Centre for Water and Waste Technology, University of New South Wales, Sydney, Australia for her contributions in the early stages of this study.

REFERENCES

Anonymous 2001 Adaptation of Dutch drinking water legislation. *Besluit van 9 januari 2001 tot wijziging van het waterleidingbesluit in verband met de richtlijn betreffende de*

kwaliteit van voor menselijke consumptie bestemd water, **31**. Staatsblad van het Koninkrijk der Nederlanden, pp. 1–53.

Bukowski, J., Korn, L. & Wartenberg, D. 1995 Correlated inputs in quanitative risk assessment: the effects of distributional shape. *Risk Analysis* **15**(2), 215–219.

De Man, J. C. 1975 The probability of most probable numbers. *European J. Appl. Microbiol.* **I**, 67–78.

Gale, P., van Dijk, P. A. H. & Stanfield, G. 1997 Drinking water treatment increases micro-organism clustering; the implications for microbial risk assessment. *J. Water Supply Res. Technol. Aqua* **46**, 117–126.

Gerba, C. P. & Haas, C. N. 1988 *Assessment of Risks Associated with Enteric Viruses in Contaminated Drinking Water 976*. ASTM Special Technical Publication, pp. 489–494.

Gibson, C. J., III, Haas, C. N. & Rose, J. B. 1999 Risk assessment of waterborne protozoa: current status and future trends. *Parasitology* **117**, 205–212.

Greenwood, M. & Yule, G. U. 1917 On the statistical interpretation of some bacteriological methods employed in water analysis. *J. Hyg.* **16**, 36–54.

Haas, C. N. 1983 Estimation of risk due to low doses of microorganisms: a comparison of alternative methodologies. *Am. J. Epidemiol.* **118**, 573–582.

Haas, C. N. 1997 Importance of distributional form in characterizing inputs to Monte Carlo risk assessment. *Risk Analysis* **17**(1), 107–113.

Haas, C. N. 1999 On modeling correlated random variables in risk assessment. *Risk Analysis* **19**(6), 1205–1214.

Haas, C. N. & Trussel, R. R. 1998 Frameworks for assessing reliability of multiple, independant barriers in potable water reuse. *Wat. Sci. Tech.* **38**(6), 1–8.

Haas, C. N., Rose, J. B. & Gerba, C. P. 1999 *Quantitative Microbial Risk Assessment*. Wiley, New York, USA.

Hijnen, W. A. M., Veenendaal, D., Van der Speld, W. M. H., Visser, A., Hoogenboezem, W. & van der Kooij, D. 2000 Enumeration of faecal indicator bacteria in large water volumes using on site membrane filtration to assess water treatment efficiency. *Water Res.* **34**(5), 1659–1665.

Hrudey, S. E. & Hrudey, E. J. 2004 *Safe Drinking Water–Lessons Learned from Recent outbreaks in Affluent nations*. IWA Publishing, London.

ILSI 1996 Risk science institute pathogen risk assessment working group a conceptual framework to assess the risks of human disease following exposure to pathogens. *Risk Analysis* **16**(6), 841–847.

Medema, G. J., Teunis, P. F. M., Havelaar, A. H. & Haas, C. N. 1996 Assessment of the dose-response relationship of *Campylobacter jejuni*. *Int. J. Food Microbiol.* **30**, 101–111.

Mons, M. N., Van der Wielen, J. M. L., Blokker, E. J. M., Sinclair, M. I., Hulshof, K. F. A. M., Dangendorf, F., Hunter, P. R. & Medema, G. J. 2007 Estimation of the consumption of cold tap water for microbiological risk assessment: an overview of studies and statistical analysis of data. *Journal of Water and Health* **5**(S1), 151–170.

Nichols, G. 2003 Using existing surveillance-based data. In *Drinking Water and Infectious Disease: Establishing the Links* (ed. P. R. Hunter, M. Waite & E. Ronchi), pp. 131–142. IWA Publishing, London, UK.

Payment, P., Berte, A., Prévost, M., Ménard, B. & Barbeau, B. 2000 Occurrence of pathogenic microorganisms in the Saint Lawrence River (Canada) and comparison of health risks for populations using it as their source of drinking water. *Can. J. Microbiol.* **46**, 565–576.

Rose, J. B., Haas, C. N. & Regli, S. 1991 Risk assessment and the control of waterborne Giardiasis. *Am. J. Public Health* **81**, 709–713.

Regli, S., Rose, J. B., Haas, C. N. & Gerba, C. P. 1991 Modeling the risk from Giardia and viruses in drinking water. *J. Am. Water Works Assoc. JAWWA* **83**(11), 76–84.

Smeets, P. W. M. H., Van Dijk, J. C., Stanfield, G., Rietveld, L. C. & Medema, G. J. 2007 How can the UK statutory Cryptosporidium monitoring be used for Quantitative Risk Assessment of Cryptosporidium in drinking water? *Journal of Water and Health* **5**(S1), 107–118.

Teunis, P. F. M. & Havelaar, A. H. 1999 *Cryptosporidium in Drinking Water. Evaluation of the ILSI/RSI Quantitative Risk Assessment Framework*. Report No. 284 550 006 RIVM, Bilthoven, The Netherlands.

Teunis, P. F. M., Havelaar, A. H. & Medema, G. J. 1994 *A Literature Survey on the Assessment of Microbial Risk for Drinking Water*. RIVM Report 734301006, RIVM Bilthoven, The Netherlands.

Teunis, P. F. M., Medema, G. J., Kruidenier, L. & Havelaar, A. H. 1997 Assessment of the risk of infection by *Cryptosporidium* or Giardia in drinking water from a surface water source. *Water Res.* **31**(6), 1333–1346.

Teunis, P. F. M., Evers, E. G. & Slob, W. 1999 Analysis of variable fractions resulting from microbial counts. *Quant. Microbiol.* **1**, 63–88.

Teunis, P., Davison, A. & Deere, D. 2004 *Short-term Fluctuations in Drinking Water Quality and their Significance for Public Health, report*. World Health Organisation, Geneva, Switzerland.

Van Gelder, P. H. A. J. M. 1999 *Statistical Methods for the Risk-Based Design of Civil Structures* Thesis Delft University of Technology, The Netherlands, available at Communications on Hydraulic and Geotechnical Engineering, ISSN:0169-6548 00-1.

Westrell, T., Bergstedt, O., Stenström, T. A. & Ashbolt, N. J. 2003 A theoretical approach to assess microbial risks due to failures in drinking water systems. *Int. J. Environ. Health Res.* **13**(2), 181–197.

WHO 2004 *Guidelines for Drinking Water Quality third edition*. WHO, Geneva.

Yao, K.-M., Habibian, M. T. & O'Melia, R. C. 1971 Water and waste water filtration: concepts and applications. *Environ. Sci. Technol.* **5**(11), 1105–1112.

First received 27 July 2007; accepted in revised form 7 October 2007. Available online March 2008