Practical applications of quantitative microbial risk assessment (QMRA) for water safety plans

P. W. M. H. Smeets, L. C. Rietveld, J. C. van Dijk and G. J. Medema

ABSTRACT

The absence of indicator organisms in drinking water does not provide sufficient guarantee for microbial safety. Therefore the water utilities are implementing water safety plans (WSP) to safeguard drinking water quality. Quantitative microbial risk assessment (QMRA) can be used to provide objective quantitative input for the WSP. This study presents several applications of treatment modelling in QMRA to answer the risk managers questions raised in the WSP. QMRA can estimate how safe the water is, how much the safety varies and how certain the estimate of safety is. This can be used in the WSP system assessment to determine whether treatment is meeting health-based targets with the required level of certainty. Quantitative data analysis showed that short events of only 8 hours per year can dominate the yearly average health risk for the consumer. QMRA also helps the design of physical and microbial monitoring. The study showed that the required monitoring frequency increases with increasing treatment efficacy. Daily monitoring can be sufficient to verify a treatment process achieving 2 log reduction of pathogens, but a process achieving 4 log reduction needs to be monitored every 15 minutes. Similarly, QMRA helps to prepare adequate corrective actions by determining the acceptable ‘down time’ of a process. For example, for a process achieving 2.5 log reduction a down time of maximum 6 hours per year is acceptable. These applications illustrate how QMRA can contribute to efficient and effective management of microbial drinking water safety.

Key words | QMRA, water safety plans

INTRODUCTION

The goal of the water safety plan (WSP) is to manage water supply such that health-based targets are met (Davison et al. 2006). To determine if the health-based targets are met, the risk of infection from the supplied water needs to be assessed. Qualitative and semi-quantitative estimates of risk have been applied in WSP, such as the risk matrix (Davison et al. 2006). The risk matrix relies on the previous experiences of those involved in the WSP development. The risk matrix requires detailed quantitative judgement on the effect of a risk such as ‘compliance impact (minor)’ or ‘public health impact (catastrophic)’. This method leads to a subjective judgement of risk that is suitable to identify ‘clear’ risks. A ‘tier one’ level of risk assessment is often applied, using log credits (indicating the level of reduction of pathogens in log units) and CT tables (where CT refers to the product of the disinfectant concentration and the contact time) to quantify treatment efficacy (Davison et al. 2006). Log credits could be effective at a screening level quantitative microbial risk assessment (QMRA) to prioritize between different sites and to identify sites that may not achieve the health-based targets (Medema et al. 2006). However, log credits do not account for site specific conditions and variations of treatment efficacy in time and the uncertainty involved (Teunis et al. 2009). In QMRA the...
risk of infection from drinking water from a treatment system is assessed by identifying the pathogen sources, monitoring pathogens in source water, describing treatment performance, contamination during distribution and consumption and applying dose-response relations (Haas et al. 1999; Smeets et al. 2008). Acceptable risk levels referred to in guidelines and legislation are very low, in the order of 1 infection per 10,000 people per year (Anonymous 2001; WHO 2004), so effective treatment is crucial for safe drinking water production. QMRA can quantify whether adequate operation and sufficient monitoring was performed to verify that the water was safe. This paper illustrates how QMRA can be used in the WSP to design microbial and non-microbial monitoring, set critical limits and prepare corrective actions based on quantitative information to achieve treatment goals.

**METHODS**

Operational treatment data was collected in the MicroRisk project (Smeets 2008). This data was used in to illustrate the preparation of corrective actions in case of chlorine dosage failure. UK statutory Cryptosporidium monitoring data was collected through the UK statutory Cryptosporidium monitoring program. All surface water treatment systems perform daily continuous monitoring for at least 23 hours, filtering 40 L/h of their produced drinking water. This results in approximately 1,000 L samples that are analysed for Cryptosporidium. The data is described in more detail in (Smeets et al. 2007). Data was presented in a complementary cumulative frequency distribution by sorting the reported concentrations, including negative samples, and attributing an equal proportion of time to each sample. By cumulating the proportions, a CCDF was plotted so that each point indicates the proportion of time that a concentration was exceeded.

The effect of short events of poor treatment on the yearly average treatment efficacy was calculated as follows. A loss of treatment efficacy due to a special event, e.g. a clogged dosing pump, was referred to as a failure. The acceptable risk criteria for failures were expressed in the form of Equation (1) where $\alpha_{acc}$ is the acceptable contribution of failures to average risk, $\pi_n$ is the nominal treatment efficacy, $\pi_f$ is the treatment efficacy during failure and $p_f$ is the proportion of time that failure occurs.

$$\alpha_{acc} \pi_n \geq \pi_f p_f$$

(1)

The probability of detecting a failure event was calculated with Equation (2). Assuming that monitoring indicated either “non-failure” or “failure” of the process, the chance of detecting a failure event $p_d$ could be determined from the chance of failure $p_f$ and the number of samples $N$:

$$p_d = 1 - (1 - p_f)^N$$

(2)

The mean treatment efficacy including failures $\pi_m$ was calculated as:

$$\pi_m = (1 - p_f) \pi_n + p_f \pi_f$$

(3)

The average yearly treatment efficacy including corrective actions was calculated with Equation (4) which is similar to Equation (3) but included three situations: nominal treatment, failure and corrective treatment.

$$\pi_m = (1 - p_f - p_c) \pi_n + p_f \pi_f + p_c \pi_c$$

(4)

where $p_c$ was the proportion of time that corrective treatment was performed and $\pi_c$ was the treatment efficacy during corrective treatment.

**RESULTS**

**Design of microbial monitoring**

Microbial monitoring before and after a treatment process, is the most direct way to assess treatment efficacy. Furthermore microbial monitoring of drinking water could also provide a direct assessment of drinking water safety. The microbial monitoring needs to provide a reliable estimate of the arithmetic mean concentration for both these applications. When organisms in water are over dispersed, high concentrations that rarely occur can still dominate the mean concentration. Monitoring will provide a range of microorganism concentrations, from which the mean concentration is calculated. A question from the risk manager could be: Have I taken enough samples to determine the mean microorganism concentration in the...
water? To answer this question, the risk manager needs to determine whether the “dominant concentrations” have been determined by monitoring. The answer to this question is illustrated through Figures 1a, b and 2. The markers show the observed *Cryptosporidium* concentrations in drinking water at three treatment sites in the UK from 2000 to 2002. The contribution of each monitored concentration to the mean was determined as the product of the concentration and the proportion of the drinking water exceeding that concentration. A low concentration of 0.01 organism/L exceeded 100% of the time resulted in the same contribution to the mean concentration as a high concentration of 1 organism/L exceeded 1% of the time. The dashed line shows combinations of proportion and concentration that contributed equally to the mean concentration. The arithmetic mean concentration at the site in Figure 1a was 0.002 oocysts/L. The observed concentrations between $2 \times 10^{-3}$ and $10^{-2}$ oocysts/L dominated this mean concentration. This is where the dashed line touches the markers of the observed concentrations. Based on this data, the average concentration was accurately determined, since higher and lower concentrations would not contribute significantly to the mean concentration. This typical shape was found for 30% of the 216 treatment sites for which data had been collected. Dominant concentrations generally occurred in less then 1%–5% of the samples in raw water, during treatment and in drinking water. So in order to generate a usable estimate of the risk, over 20 to 100 microbial samples at any point in the system would typically be required.

A complicating factor is the occurrence of ‘special events’. Figure 1a shows the concentrations due to normal variations, so high concentrations can be considered ‘normal events’. Figure 1b shows a ‘special event’ where very high concentrations occurred e.g. due to human error or equipment failure causing a curve break. In Figure 1b the dashed line touches the markers at the highest monitored concentration. The concentration during the special event therefore dominated the average concentration and thus the yearly average risk at this site. These special events were observed in 10% of the datasets from 216 treatment sites in the UK. At the example site in Figure 1b, other monitoring efforts, apart from microbial monitoring, would be required to verify that these special events did not occur even for a few hours per year (a proportion of $10^{-3}$ translates into eight hours per year). Methods for event monitoring will be discussed later.

A second complicating factor is the detection limit or sample volume. At treatment sites with lower concentrations of *Cryptosporidium* in the treated water, the observed concentrations looked like Figure 2 which was typical for 60% of the monitored UK sites. The arithmetic mean concentration was $6 \times 10^{-5}$ oocysts/L based on the monitoring results. This mean concentration was dominated by the lowest observed concentration (at the detection limit) of $7 \times 10^{-4}$ oocysts/L exceeded 7% of the time. However, an intuitive interpolation to lower concentrations would exceed the dashed line, thus dominating the mean concentration and
therefore the average risk. Larger volumes taken at a lower frequency could be used to determine the dominating concentration. Stochastic modelling of the treatment as performed in QMRA could provide an estimate of these low concentrations (Smeets et al. 2008), based on raw water concentrations and removal by treatment.

Design of process monitoring

When the treatment assessment in the WSP indicates that the system is capable of producing safe water, the system needs to be monitored to verify that treatment targets were met. Therefore one of the monitoring goals is to verify that no special events have occurred that could have a significant impact on the assessed risk (Figure 1b). Since this often requires very frequent monitoring, microbial monitoring is generally not feasible. However, other parameters that may be monitored on-line can be used to verify the system was working as expected. Equipment monitoring such as dosing pump flow can also be used to verify that the system was operational, or to detect moments of failure. Deviation of any of these parameters could lead to reduced treatment efficacy, which may be hard to quantify. A conservative approach is to assume complete failure in case of a deviance. The following examples assume that monitoring either indicates compliance or failure (no pathogen reduction), however if efficacy during failure could be quantified in more detail this could be incorporated in the calculations.

First the acceptable level of increase of risk due to failure $a_{acc}$ in Equation (1) needs to be chosen. The risk due to the event will add to the nominal risk. Risks due to special events might result in a higher average risk than the nominal risk, which would not be considered acceptable. In this example the risk from failures is considered acceptable when on average it equals the nominal risk, so $a_{acc}$ equals 1. A more strict condition could be set by choosing a lower value for $a_{acc}$. Complete failure of the treatment process during an event was assumed in this example, so treatment efficacy during failure $\pi_f$ equalled 1. In that case the acceptable proportion of time that failure occurs $p_f$ equalled nominal reduction $\pi_n$. The probability of detecting a failure event was calculated with Equation (2) assuming that monitoring indicated either “non-failure” or “failure” of the process. By choosing the probability of detection $p_d$ at the desired confidence level (e.g. 90%, 95%, 99%) the required number of measurements $N$ could be calculated.

The safety of the system is verified when no failures are detected with $N$ measurements over the assessed period. Table 1 shows the required number of measurements in relation to the nominal log reduction to verify at the 95% confidence level that failure events did not significantly increase the risk. The monitoring frequency for the assessment of a one year period was also calculated. Approximately 30% less measurements are required for a 90% confidence level, respectively 50% more measurements are required for a 99% confidence level.

Table 1 | Number of required monitoring records to verify at the 95% confidence level that failure events do not significantly add to the risk when compared to nominal reduction

<table>
<thead>
<tr>
<th>Nominal log reduction</th>
<th>#/year</th>
<th>Monitoring interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1</td>
<td>1 year</td>
</tr>
<tr>
<td>1</td>
<td>30</td>
<td>1 week</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>1 day</td>
</tr>
<tr>
<td>3</td>
<td>3,000</td>
<td>3 hours</td>
</tr>
<tr>
<td>4</td>
<td>30,000</td>
<td>15 min</td>
</tr>
<tr>
<td>5</td>
<td>300,000</td>
<td>2 min</td>
</tr>
<tr>
<td>6</td>
<td>3,000,000</td>
<td>10 sec</td>
</tr>
<tr>
<td>7</td>
<td>30,000,000</td>
<td>1 sec</td>
</tr>
</tbody>
</table>
Table 1 shows that with daily sampling, no more than 2 log reduction can be verified. Above that, more frequent monitoring is required, leading to on-line monitoring. Some treatment plants claim 7 log inactivation of viruses with a disinfection system, so at a 100,000 m$^3$/d plant every 3 litres must be monitored to be 95% confident that all water was sufficiently treated. If the criteria are set stricter, for example setting $\alpha_{acc}$ to 10%, ten times as many samples are required.

A multiple barrier system is easier to monitor. When 6 log inactivation is achieved through 1 barrier (treatment process) 3 million measurements per year are required. The continuous measurement of UV radiation intensity is an example of such intensive monitoring to verifying absence of failure of the highly effective UV disinfection process. When three barriers of 2 logs are placed in series, only 500 measurements per year are required for each process, resulting in 900 measurements in total. So daily monitoring of each process step is then sufficient. This approach assumes that treatment process failures are independent and that failure of the first step does not automatically coincide with failure of the second step. Risk managers can minimise the risk of such consecutive failures in their design of the process for example by separating power supplies and placing physical barriers between equipment so that physical hazards (e.g. flooding or fire) would not affect both processes.

### Setting critical limits

A treatment system can be designed to provide exactly the right level of treatment to meet the health-based targets. However, in practice the risk manager needs to account for variations and inaccuracies in order to run a practical and stable process. The simulated example in Figure 3 is used to illustrate how QMRA can be used to set critical limits and setpoints. The health-based treatment target for the disinfection process simulated in Figure 3 was 3 log inactivation of viruses. In this example the temperature and flow conditions were constant for the whole period. The required chlorine residual of 1.7 mg/l for 3 log inactivation was determined by disinfection modelling. So if the process was constantly run at exactly this concentration, the yearly health-based target would be met. This level is therefore considered the critical limit. However, in practice the chlorine dose would vary due to operational variation. Chlorine dosing was controlled by an automatic control loop in this example. The chlorine residual was measured and the dosing rate was adjusted if the measured concentration deviated from the setpoint. As a result, the chlorine residual level would typically vary between the operational limits if the system was working within specification, as indicated in Figure 3.

At hour 20 the chlorine dosing pump got clogged, resulting in a chlorine level below the lower operational limit. This triggered an alarm and the operator was able to clean the pump and restore normal operation. This event did not affect the average treatment performance in a way that the health-based target would not be met, since the critical limit was not exceeded. At hour 80 the chlorine dosing pump failed again, however this time the operator was not in time to restore the system before the critical limit was exceeded. The short period of time that disinfection was ineffective reduced the average treatment efficacy since the critical limit was exceeded. Therefore the operator needed to take corrective actions, such as starting emergency chlorination of the distributed water, in order to achieve the health-based target over the total period.

This example illustrated that setting critical limits was not merely determining the lowest chlorine dose to meet the target, but also included setting operational setpoints and operational limits. Quantifying limits and setpoints is related to many site specific issues such as the treatment target, the normal variability of the process, the probability...
of events, the response time of the operator and the effect of corrective actions. Simply applying a higher setpoint for chlorine dose can have adverse effects in other objectives (cost, Disinfectant By-Product (DBP) formation, maintenance/operation) and therefore is not an option. QMRA can provide decision support for choosing setpoints and limits when some realistic assumptions are made. The mean treatment efficacy over a period was calculated with Equation (3). The proportion of time that the system fails $p_t$ and treatment efficacy during failure $\pi_t$ can be roughly quantified by means of the risk matrix in the WSP. In the example it was assumed that operator response could take up to 8 hours to restore the system to normal operation, an event could occur once a year and disinfection would fail completely during an event. Given these assumptions, $p_t$ was $8/(365 \times 24)$, $\pi_t$ equaled 1 and $\pi_m$ equaled 0.001 ($3 \log$ reduction). Using Equation (3) the nominal treatment efficacy $\pi_n$ to achieve $\pi_m$ was 0.000085 ($4.1 \log$ reduction). Disinfection modelling showed that a chlorine residual of 2 mg/l was required to achieve $4.1 \log$ reduction, therefore the lower operational limit was set to 2 mg/l. Due to normal variation in the process, observed from hour 25 to 75 in Figure 3, a setpoint of 2.3 mg/l was needed to maintain this limit. This example illustrated the basic approach to setting limits and setpoints. The method could be adapted to include other variables such as temperature and flow variations and partial treatment failure.

One of the objectives of the multiple barrier system is to balance out peaks or failure at one point by adequate treatment at another point. Total pathogen reduction at a surface water treatment site is often many orders of magnitude due to a large number of treatment steps. The total treatment can be optimized such that the combination of reduction by individual treatment steps combined provides exactly the required reduction to meet the health-based targets. This could be extended to an on-line control tool for pathogen reduction to respond to changes in process conditions. The disinfection model could be used in an algorithm programmed into an automatic operation system to constantly adjust the critical limits, operational limits and setpoint for chlorine dosing to maintain the target level of inactivation. The treatment would thus be controlled by the combined effect of process conditions rather than control of individual parameters.

All treatment steps were considered to be independent in this study. In practice some interaction between treatment steps can be expected. Positive interaction occurs when failure at one step leads to failure of a consecutive step(s). For example, failure of pre-oxidation leads directly to less inactivation, but also increases oxidant demand at a later disinfection step, thereby reducing its efficacy. Negative interaction can occur when failure at an early step leads to increased performance of a consecutive step, e.g. poor sedimentation increases particle load to the filters which then work more effectively. Although these mechanisms have been observed, current knowledge is insufficient to create a deterministic model that describes these relations.

Preparing corrective actions

Even at a well managed system failures can occur that compromise drinking water safety. Such events could be detected by monitoring. A temporal decrease of drinking water quality could have a significant effect on the yearly average risk. When critical limits are exceeded, corrective actions need to bring back water quality to an acceptable level to comply with the yearly risk of infection. This is illustrated with the example of chlorine inactivation of viruses in Figure 4. The target inactivation by this system was 2.5 logs, which was generally achieved through process control. The risk manager would need to prepare a plan for corrective actions if failure occurred. Shutting down treatment completely in case of failure was not an option since this would disrupt the other processes and the distribution system would loose pressure thus risking...
Ingress of contaminated water into the distribution system. An emergency UV disinfection unit which could achieve 4.5 log inactivation of viruses was considered as a corrective measure. The first question for the risk manager was how quickly the emergency equipment needed to be in place. Or, stated differently: how long could this failure be allowed to continue without compromising the treatment target? QMRA could be used to answer this question. Figure 4 shows the calculated virus inactivation based on monitored chlorine residual concentration. During normal operation from 27 April to 19 May, generally 2.5 to 3 log inactivation was achieved (black line), resulting in a running average inactivation that complies with the target of 2.5 log inactivation. On 20 May the chlorine dosing failed completely (as an example), resulting in no disinfection. The grey line shows the level of inactivation that could be achieved with the emergency equipment. The average yearly treatment efficacy was calculated for different response times with Equation (4). In this example \( \pi_m \) and \( \pi_n \) were 0.0032 (2.5 log reduction), \( \pi_c \) was 0.000032 (4.5 log) \( p_n \) was \( \frac{139}{365} \) (days until 20 May), \( p_c = 1 - p_n - p_f \) and \( p_f \) varied between 0 hours and 17 days.

The dashed line in Figure 4 shows the achievable yearly average inactivation in relation to the time the corrective action is started. It was clear from Figure 4 that the achievable yearly average quickly decreased in time. In this example, corrective measures needed to be taken within six and a half hours to comply with the yearly target of 2.5 log reduction. Risk managers can use similar calculations for decision support on emergency measures and emergency procedures in relation to the health-based treatment target, the critical limits of the treatment process and the achievable corrective action.

CONCLUSIONS

Water utilities have started to use water safety plans (WSP) to assess and improve the safety of the produced drinking water. Resources for assessments and improvements are limited and therefore need to be used effectively and efficiently to provide the most benefit for health. Quantitative microbial risk assessment can be used to quantify several questions that are raised in the WSP. Without QMRA the system assessment in the WSP relies on risk manager experience and log-credits from industry standards to quantify drinking water safety. QMRA does not only tell us how safe the water is, but also how much the safety varies and how certain we are that we are meeting health-based targets. The QMRA methods presented in this study can be used to quantify and reduce uncertainties by using site specific full-scale information. QMRA can be used in the system assessment and provides decision support on other issues of the WSP. This study provided examples of QMRA applications to design physical and microbial monitoring, to set critical limits, and to prepare corrective actions. Thus QMRA can contribute to efficient and effective management of microbial drinking water safety.

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.

REFERENCES


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

