Legionnaires’ disease: evaluation of a quantitative microbial risk assessment model

Thomas W. Armstrong and Charles N. Haas

ABSTRACT

Background: The quantities of Legionella vary considerably from natural waters to water in contaminated domestic hot water supplies, whirlpool spas and cooling towers, with the risk for LD rising as the Legionella counts grow. We currently report the results from our Quantitative Microbial Risk Assessment (QMRA) model evaluation. We developed the LD QMRA model to better understand Legionella exposure risks.

Methods: Using an animal data derived model for LD, we calculated risks from estimated exposures for a whirlpool spa outbreak, two hot spring spa outbreaks and compared the results to the reported LD risks.

Results: The QMRA model shows agreement (generally less than an order of magnitude discrepancy) with the reported Legionnaires’ disease sub-clinical severity infection, clinical severity infection, and mortality risks.

Conclusions: The LD QMRA model may lead to risk based limits to supplement the current guidance on Legionella control in cooling towers, whirlpool spas and other potential exposure sources. The verification of QMRA for LD also suggests the techniques, given suitable animal model data, may be useful in quantifying human response to other airborne pathogens.

Key words: Legionnaires’ disease, Legionella, Legionella pneumophila, risk assessment, quantitative microbial risk assessment

ABBREVIATIONS AND TERMINOLOGY

LD Legionnaires’ disease.
QMRA Quantitative microbial risk assessment.
LD\textsubscript{50\%} Lethal dose 50\%, or the dose causing 50\% mortality.
ID\textsubscript{50\%} Infectious dose 50\%, or the dose causing a 50\% infection rate.
CFU Colony forming units, where each colony on a microbial culture plate presumably represents a single organism from the initial inoculation.

Clinical infection, clinical severity infection LD infection illness of severity requiring supportive medical care, including hospitalization and treatment for pneumonia.
Sub-clinical infection, sub-clinical severity infection LD infection which did not progress to pneumonia requiring medical treatment. This milder infection may or may not have led to seroconversion.

INTRODUCTION

Legionnaire’s disease (LD), reported as caused by the bacterium Legionella pneumophila (Lp) in 1977 (McDade et al. 1977), remains at 30 years later a significant cause of community acquired and nosocomial pneumonia, although
the actual incidence rates are uncertain (Millar 1997; Fields et al. 2002). Legionella are widespread, with numerous species and strains found in soils and water supplies throughout the world, and LD is widespread as well (Fliermans et al. 1981; Joly et al. 1984; Korvick et al. 1987; Ortiz-Roque & Hazen 1987; Steele et al. 1990; Zacheus & Martikainen 1994; Koide et al. 2001; Riffard et al. 2001). Given suitable temperature and nutrient conditions, Legionella can amplify in cooling towers, hot water distribution systems, and other water-containing systems. For cooling towers, the basis for the guidance which exists for Legionella control limits is not clearly documented. The guidance apparently evolved from experience, and may largely derive from a study of the counts of Legionella in two cooling towers tied to LD outbreaks versus counts of the bacteria in other cooling towers not associated with outbreaks (Shelton et al. 1994). The basis for any limits for whirlpool spas, hot water distribution systems, and other potential sources of Legionella exposure are not well described either. Given the seemingly limited basis for the Legionella control guidance, we undertook research on another approach for quantitatively evaluating the risks from Legionella exposure.

In 1983, Haas reported on models to estimate risks from low dose exposure to waterborne bacteria and viruses (Haas 1985). Since then the QMRA techniques have been applied in many other water and food microbial risk assessments, but less commonly for airborne microbial exposures and risks. Although many investigators reported animal dose-response data for Legionella since 1976, prior risk assessment reports are generally limited (O’Brien & Bhopal 1995) or qualitative and our research appears to be the first to use the available data and QMRA techniques to model human LD risks from aerosol exposure to the organism. Our QMRA development for LD progressed in three parts. Figure 1 provides a summary of the relationships of the three parts of our QMRA development project for Legionella. Part 1 involved a search for published animal model reports with dose and response data following controlled inhalation dosing with aerosol containing viable and virulent Legionella (primarily L. pneumophila serogroup 1). The diversity of animal models used, and the quantitative diversity of their responses led to further consideration of which of the available animal models could serve as the best basis for extrapolation to humans. Guinea pigs appeared most suitable based on a number of factors, but especially due to in vitro data showing quantitative similarities for Legionella uptake, survival and replication in that species and human macrophages (Rechnitzer et al. 1992). Legionella produce LD following intracellular replication in alveolar macrophages, so similar macrophage uptake and replication is a particularly important model selection aspect. In vivo and in vitro macrophage data showed most mouse strains were resistant to Legionella infection (Izu et al. 1999; Wright et al. 2003). Data for non-human primates were limited (Kishimoto et al. 1979; Baskerville et al. 1983; Jacobs et al. 1984; Dowling et al. 1992), but also indicated resistance to Legionella infection compared to other susceptible species. Further details on other comparative immune response factors considered and full references on the animal model selection issues are available in a dissertation and conference poster (Armstrong & Haas 2003; Armstrong 2005).

Following the selection of the Guinea pig model, data for mortality (Baskerville et al. 1981; Fitzgeorge et al. 1985) and non-lethal infection (Muller et al. 1985) were both used with dose-response models. The models evaluated were the exponential, approximate beta-Poisson, exact beta-Poisson, Weibull, logistic and probit. The dose-response modeling work was primarily done in Excel® using Solver for maximum-likelihood optimization of the model parameters (Haas 1994), with Mathematica® used to a lesser extent. Each model’s fit to the data was evaluated via $\chi^2$ techniques (Haas 1994). Due to their mechanistic basis (Haas et al. 1992; Reinhold et al. 1994; Kishimoto et al. 1983; Rechnitzer et al. 1992; Koide et al. 2001; Wright et al. 2003).
the enrichment of bacteria of multiple species in aerosol spa’s injected air was considered. Many reports demonstrate whirlpool spa assessment, active aerosol generation by the information on the underlying exposure factors. For the ing circumstances, and also differed in the extent of different approaches since the underlying events had differ-

assessments for the three spa outbreaks employed two

human risk would be available, and our QMRA research such data had been available, direct quantitative evidence on cases’ exposures as well as the incidence rates of LD. If reports available which provided quantitative data on LD the QMRA model’s adequacy. There were no LD outbreak exposure and LD rate information available for evaluating

models provided an improved fit to the data. The dose-

response models were then used to project human risk at doses below the animal experimental dose ranges.

The guinea pig LD_{50%} experiments allowed the infections to run its course to mortality or recovery, without supportive care intervention with antibiotics or other methods. On that basis alone, the guinea pig data may adequately predict severe morbidity in humans, but may project higher mortality than would occur in humans who then receive medical intervention. Part 1 of the project, the dose-response modeling and extrapolation to humans, comprises the actual LD QMRA model. Dose scaling was not used, for mechanistic considerations of microbial uptake and replication. No adjustments were applied for interspecies relative susceptibility given the similarities in guinea pig and human alveolar macrophage response to Legionella infection. Manuscripts on Part 1 includes additional discussion of the above aspects (Armstrong 2005; Armstrong & Haas 2007b).

Part 2 covered the methods and results for the assessment of Legionella exposures for three relatively well documented spa-related outbreaks of LD. Only a brief summary fits in the scope of this current report. Additional details on Part 2 are available elsewhere (Armstrong 2005; Armstrong & Haas 2007a). To a large degree, the LD outbreak exposure assessment work was motivated by the need to develop exposure information for outbreaks where the LD disease rate was well known, and thus have both exposure and LD rate information available for evaluating the QMRA model’s adequacy. There were no LD outbreak reports available which provided quantitative data on LD cases’ exposures as well as the incidence rates of LD. If such data had been available, direct quantitative evidence on human risk would be available, and our QMRA research would perhaps not have been undertaken. Our exposure assessments for the three spa outbreaks employed two different approaches since the underlying events had differing circumstances, and also differed in the extent of information on the underlying exposure factors. For the whirlpool spa assessment, active aerosol generation by the spa’s injected air was considered. Many reports demonstrate the enrichment of bacteria of multiple species in aerosol compared to the bulk water content (Blanchard & Syzdek 1970; Quinn et al. 1975; Hejkal et al. 1980; Blanchard & Syzdek 1982; Blanchard 1983; Colbourne et al. 1987; Blanchard 1989; Georgescu et al. 2002). This potential enrichment was an important consideration in the whirlpool spa exposure assessment. The other two spa outbreaks were for natural thermal spring spas, without active aeration, and a bacterial water-to-air partitioning coefficient was developed and used to estimate exposures.

In this current report, we describe Part 3 of our Legionnaires QMRA work. That is the evaluation of our QMRA model’s predictive adequacy for LD. An extended abstract summarizes the methods and evaluation results (Armstrong & Haas 2006) and a doctoral dissertation covers the whole of the research (Armstrong 2005).

Infection with Legionella (dominantly Legionella pneumo-
phil a) in humans produces a range of outcomes, possibly due to variable ranges of exposure, variations in human susceptibility, and variations in the virulence of the Legionella organism with the strains or growth conditions. These infection outcomes include:

- Pontiac Fever (Kaufmann et al. 1981), a febrile illness with no associated mortality
- Apparently silent (or sub-clinical severity) Legionella infections (Lattimer et al. 1979; Boshuizen et al. 2001)
- Severe clinical manifestations of Legionnaires’ disease (Edelstein & Meyer 1984) including pneumonia and typically, for a fraction of those ill with LD, mortality. This “clinical severity” stage of LD would generally involve hospitalization and supportive medical care.

Given the different disease stages possible, we used different guinea pig exposure reports (Baskerville et al. 1981; Muller et al. 1983) for the different LD stages. We did not include Pontiac Fever. We modeled the sub-clinical infection rates for humans using data for rates of infection in guinea pigs which recovered (with zero mortality in any dose group) following a fever induced by aerosol exposure to L. pneumophila (Muller et al. 1983). We used data for the inhalation LD_{50%} in guinea pigs (Baskerville et al. 1981; Fitzgeorge et al. 1983) to model the rates of clinical severity infection and mortality in humans. After initial examination of a wider range of models, we used the exponential and beta-Poisson models for subsequent work (Armstrong 2005;
Armstrong & Haas 2007b). The exponential and beta-Poisson models gave similar results. For the exponential model (Equation 1), the best-fit results, using the data reported by Muller et al. 1983, suggest that the subclinical LD rate for dose d is given with the parameter $r = 0.06$.

$$P_1(d) = 1 - e^{(-r \cdot d)}$$ (1)

Where $P_1(d)$ = the risk at dose d of *Legionella*, $r$ = the model parameter.

For the best-fit model using the data of Baskerville et al. 1981, the risk of clinical severity LD (or mortality) from dose d is predicted with the model parameter $r = 1.07 \times 10^{-4}$. With the LD QMRA model, for a given level of exposure, the expected risk may then be calculated, or in reverse, for a given level of risk, the dose may be estimated.

Published reports provided the rates for sub-clinical and clinical-severity infection for exhibitors who worked during a large LD outbreak in the Netherlands in 1999 (Boshuizen et al. 2000; Boshuizen et al. 2001; den Boer et al. 2002). A more recent report on the minimum infection rates for the exhibitor staff (Boshuizen et al. 2006) supplements the earlier data. The newer estimates account for a possible bias in the prior post-outbreak seroconversion estimates by considering possible prior-to-outbreak antibody levels.

Published reports also provided data on rates of clinical infection and mortality among the visitors to natural hot spring spas for several LD outbreaks in Japan (Sugiyama et al. 2000; Anonymous 2003; Yabuuchi & Agata 2004; Okada et al. 2005). QMRA model calculated rates of infection and comparisons to the LD outbreaks’ reported infection rates follow.

**METHODS FOR LD QMRA MODEL EVALUATION**

**Whirlpool spa outbreak, reported infection rates**

The reported incidence of sub-clinical severity *Legionella* infections amongst the exhibitor staff during the exposure window of the Legionnaires’ disease epidemic was 80% (95% CI 75.6 to 82.6%) (Nagelkerke et al. 2003), with more recent updates (Boshuizen et al. 2006) on minimal infection rates of 40 and 13% for the more exposed and lesser exposed groups, respectively. Only a fraction of these workers with sub-clinical severity infections progressed to clinical severity LD. The more exposed group worked primarily within 15 metres of the putative *Legionella* source, and the lesser exposed group worked primarily at more than 15 metres from the infection source. The earlier reported data includes the number of workers who were at various distances and the respective infection rates based on seroconversion prevalence. The more recent estimated minimum infection rates consider possible prior antibody levels from past exposure as potentially conferring resistance to infection during the outbreak. Based on the 2003 report, the estimated rate of infection for those exhibitors who worked primarily within 15 metres of the likely *Legionella* aerosol source was 85%; for exhibitors at 15 to 50 metres 78.7%; at 30 to 60 metres 44.5% and finally, for those at 60 to 90 metres distance, the infection rate was 25.9%. The weighted average sub-clinical infection rate for the groups at greater than 15 metres is 38%. Table 1 summarizes these rates and the calculated exact binomial confidence intervals for the < 15 metre and composite > 15 metre groups. These sub-clinical infection rates are as reported with the serum bank rates for control. A figure in the investigation report (Nagelkerke et al. 2003) suggests a sub-clinical severity infection rate of 90 to 100% at less than 15 metres, but that report’s table of rates versus distance does not include these data. Table 1 also includes the more recently estimated (Boshuizen et al. 2006) minimum sub-clinical infection rates and the confidence interval on those minimum rates, at 40% (34- -48) for the more exposed group and 13% (8 to 18) for the lower exposed group.

Table 1 also shows the clinical severity infection rates reported (Boshuizen et al. 2000) and calculated binomial confidence intervals for the < 15 metre and composite > 15 metre groups. The outbreak reports referenced included the confidence intervals on the infection rates. The recent minimum infection rate estimates (Boshuizen et al. 2006) do not affect these clinical severity LD infection rates.

**Risk calculations – whirlpool spa outbreak, estimates of infection rates**

Following the work on animal model selection, interspecies considerations and the dose-response models summarized above, we calculated probability distributions
for the estimated risks of a) sub-clinical infection and b) clinical severity infection or mortality. The calculations employed Monte Carlo techniques in Crystal Ball with an Excel spreadsheet. Our dose-response models used the data for the sub-clinical infection (Muller et al. 1983) and for the clinical severity infection and mortality (Baskerville et al. 1981).

For the sub-clinical infection and mortality models, we calculated risk estimates using the dose-response modeling approaches summarized above. The resulting rates are as exposure dose-response distributions. Table 3 summarizes the estimated Legionella dose distribution, as well as the mean, median, 25th and 75th percentiles, and the 95 percent interval (2.5 to 97.5th percentiles) of the resulting calculated risk distributions.

Our statistical criterion for reasonable model validity was partial overlap of the confidence intervals of the calculated risks and the risks reported in the LD outbreak investigations.

Natural hot spring spa outbreaks, reported data on infection rates

An outbreak of Legionnaires’ disease in Miyazaki Prefecture, Japan in July 2002 (Anonymous 2003; Yabuuchi & Agata 2004; Okada et al. 2005) involved 295 persons (109 were hospitalized, 34 were confirmed as LD by sputum culture and 7 died). The reported Legionella content of the spa water was $1.5 \times 10^7$ CFU per litre. Among the 19,773 persons who visited the spa between 20 June and 23 July of 2002, 295 became ill with multiple symptoms consistent with Legionnaires’ disease. Of the suspect cases, the investigative team diagnosed 34 as LD by sputum culture, urinary antigen or serum antibody titer. However, for the numerator for the clinical severity infection rate for this outbreak, we used the 109 hospitalized suspected cases. Generally, LD outbreak reports do not have definitive diagnostic test results for all of the included cases since the clinical emphasis is on treatment and that goal may not always require definitive diagnostic tests. It is not certain that all of the reported spa visitors had opportunity for exposure to contaminated aerosols, but no alternative to the assumption that all had exposure opportunity seems reasonable. Possibly, a number of the persons had multiple visits to the spa, but this cannot be determined either. Using the number of hospitalized cases and the symptoms as sufficient evidence of infection, we calculated the reported clinical infection rate as $109/19,773$ or $5.5 \times 10^{-3}$ (95% confidence interval 4.6 to $6.6 \times 10^{-3}$) and the reported mortality rate as $7/19,773$ or $3.5 \times 10^{-4}$ (95% confidence interval 1.6 to $7.2 \times 10^{-4}$). Table 2 shows these rates and their calculated exact binomial or Poisson confidence intervals. These confidence intervals (and other Poison or binomial confidence intervals calculated for this report) were calculated via a web applet (Anonymous 2007) following initial hand calculation and verification of the applet results. The confidence intervals for these data are essentially the same (differing in only the second significant digit) from either the exact binomial or Poisson confidence interval calculations.

Presuming a similar ratio of sub-clinical to clinical infections as from the West Frisian outbreak (clinical cases were 1.5 to 5% of the sub-clinical cases), then the estimated

Table 1 | West Frisian outbreak reported sub-clinical infection rates and calculated 95% confidence intervals

<table>
<thead>
<tr>
<th>Distance</th>
<th>Reported subclinical rate % (95% CI)</th>
<th>Minimum subclinical rate % (95% CI)</th>
<th>Reported clinical-severity infection rate % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 m and less</td>
<td>85 (72–92)</td>
<td>40 (32–48)</td>
<td>4.3 (1.4–9.9)</td>
</tr>
<tr>
<td>15 to 30 m</td>
<td>78.7</td>
<td>–</td>
<td>3 (NA)</td>
</tr>
<tr>
<td>30 to 60 m</td>
<td>44.5</td>
<td>–</td>
<td>1.9 (NA)</td>
</tr>
<tr>
<td>60 to 90 m</td>
<td>25.9</td>
<td>–</td>
<td>0.4 (NA)</td>
</tr>
<tr>
<td>&gt;15 m (weighted average)†</td>
<td>38 (32–43)</td>
<td>13 (8–18)</td>
<td>1.1 (0.36–2.6)</td>
</tr>
</tbody>
</table>

†Calculated by weighting with the reported number workers in each zone.

Data source citations: (Boshuizen et al. 2000; Boshuizen et al. 2001; den Boer et al. 2002; Boshuizen et al. 2006).
Table 2 | Miyazaki and Shizuoka prefecture outbreaks reported infection rates and calculated 95% confidence intervals

<table>
<thead>
<tr>
<th></th>
<th>Miyazaki reported rate (95% CI)</th>
<th>Shizuoka reported rate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>$3.5 \times 10^{-4}$ (1.6–7.2 $\times 10^{-4}$)†</td>
<td>$3.5 \times 10^{-5}$ (6.3 $\times 10^{-6}$–1.3 $\times 10^{-4}$)†</td>
</tr>
<tr>
<td>Clinical infection</td>
<td>$5.5 \times 10^{-3}$ (4.6–6.6 $\times 10^{-3}$)†</td>
<td>$4 \times 10^{-4}$ (2.6–6.1 $\times 10^{-4}$)†</td>
</tr>
<tr>
<td>Projected sub-clinical infection</td>
<td>0.1–0.4 (0.09–0.4)‡</td>
<td>$8 \times 10^{-7}$–3 $\times 10^{-2}$ (7 $\times 10^{-3}$–3 $\times 10^{-2}$)‡</td>
</tr>
</tbody>
</table>

†Estimated using ratio of to sub-clinical to clinical severity rates reported for the West Frisian Outbreak.
‡Binomial confidence interval.
severity infection risks with the reported clinical severity risks and estimated mortality risks. These tables and figures include results from both the beta-Poisson and Exponential models. Both models consistently give results identical to two significant figures. The exposure estimates shown are those based on the alternative assumptions of no aerosol bacterial enrichment and a factor of 10 enrichment in the aerosol compared to the bulk water content. The basis for this is discussed in the Introduction section, with the aerosol enrichment considered the more likely scenario.

### Table 3: Calculated risks for the West Frisian outbreak worker groups

<table>
<thead>
<tr>
<th>Percentile of the risk</th>
<th>10 fold aerosol enrichment</th>
<th>No enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated dose</td>
<td>B-poisson risk</td>
</tr>
<tr>
<td><strong>Subclinical infection risk</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 15 metres from the Whirlpool Spa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7.5 CFU</td>
<td>$3.5 \times 10^{-1}$</td>
</tr>
<tr>
<td>Mean</td>
<td>10 CFU</td>
<td>$3.9 \times 10^{-1}$</td>
</tr>
<tr>
<td>2.5%</td>
<td>1.3 CFU</td>
<td>$8.4 \times 10^{-2}$</td>
</tr>
<tr>
<td>25.0%</td>
<td>2.2 $\times 10^{-1}$</td>
<td>$2.2 \times 10^{-1}$</td>
</tr>
<tr>
<td>75.0%</td>
<td>5.3 $\times 10^{-1}$</td>
<td>$5.3 \times 10^{-1}$</td>
</tr>
<tr>
<td>97.5%</td>
<td>34.4 CFU</td>
<td>$8.8 \times 10^{-1}$</td>
</tr>
<tr>
<td>Greater than 15 metres from the Whirlpool Spa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5.9 CFU</td>
<td>$2.9 \times 10^{-1}$</td>
</tr>
<tr>
<td>Mean</td>
<td>6.8 CFU</td>
<td>$3.2 \times 10^{-1}$</td>
</tr>
<tr>
<td>2.5%</td>
<td>1.4 CFU</td>
<td>$9.1 \times 10^{-2}$</td>
</tr>
<tr>
<td>25.0%</td>
<td>2.0 $\times 10^{-1}$</td>
<td>$2.0 \times 10^{-1}$</td>
</tr>
<tr>
<td>75.0%</td>
<td>4.1 $\times 10^{-1}$</td>
<td>$4.1 \times 10^{-1}$</td>
</tr>
<tr>
<td>97.5%</td>
<td>18.5 CFU</td>
<td>$7.0 \times 10^{-1}$</td>
</tr>
<tr>
<td><strong>Clinical severity infection risk</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 15 metres from the Whirlpool Spa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7.5 CFU</td>
<td>$6.4 \times 10^{-4}$</td>
</tr>
<tr>
<td>Mean</td>
<td>10 CFU</td>
<td>$8.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>2.5%</td>
<td>1.3 CFU</td>
<td>$1.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>25.0%</td>
<td>3.7 $\times 10^{-4}$</td>
<td>$3.7 \times 10^{-4}$</td>
</tr>
<tr>
<td>75.0%</td>
<td>1.1 $\times 10^{-3}$</td>
<td>$1.1 \times 10^{-3}$</td>
</tr>
<tr>
<td>97.5%</td>
<td>34.4 CFU</td>
<td>$3.2 \times 10^{-3}$</td>
</tr>
<tr>
<td>Greater than 15 metres from the Whirlpool Spa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5.9 CFU</td>
<td>$4.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>Mean</td>
<td>6.8 CFU</td>
<td>$6.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>2.5%</td>
<td>1.4 CFU</td>
<td>$1.4 \times 10^{-4}$</td>
</tr>
<tr>
<td>25.0%</td>
<td>3.2 $\times 10^{-4}$</td>
<td>$3.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>75.0%</td>
<td>7.6 $\times 10^{-4}$</td>
<td>$7.6 \times 10^{-4}$</td>
</tr>
<tr>
<td>97.5%</td>
<td>18.5 CFU</td>
<td>$1.7 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
There were no reported fatalities in the exhibitor groups, possibly due to either chance, particularly successful clinical care, or a lower prevalence of risk factors amongst the exhibition workers than amongst those typically succumbing to Legionnaires’ disease. Had the experience been more typical, given the clinical case rate and an assumed 5 to 15% fatality rate amongst those clinically ill as the generally recognized mortality rate range in Legionnaires’ disease outbreaks (Schaechter et al. 1999), we estimated a mortality rate.

From Figures 2–5, it appears possible that the assumption of no aerosol enrichment of Legionella compared to the bulk water content yields underestimates of risk. The aerosol enrichment assumption yields improved estimates compared to the reported rates. We thus dropped the no enrichment scenario and carried forward the 10 fold
enrichment results in subsequent model comparisons. We based this decision on the better agreement with the reported risks, and also on the credibility of aerosol bacterial enrichment given the literature on the topic, as cited in the Introduction section above.

The reported rates of sub-clinical infection fall into the range of the calculated risk distributions, as shown in Table 4 and Figures 2 and 3. The mean reported sub-clinical rate (Nagelkerke et al. 2003) for the group at less than 15 metres, based on seroprevalence falls at the 96th percentile of the calculated risk distribution, and the lower and upper confidence limits on the reported rate fall at the 90th and 98th percentiles, respectively. Using the minimum rates which consider the possible prior antibody levels (Boshuizen et al. 2006), the overlap is considerably greater, with the mean at the 58th percentile of the calculated distribution, with the confidence intervals from 45 to 69% and the calculated distribution thus well encompasses the reported mean.

The reported sub-clinical infection rate (Nagelkerke et al. 2003) for the worker group at greater than 15 metres aligns well also, with the mean at the 70th percentile, and the upper and lower confidence limits on the reported rate falling at the 60th to 79th percentile of the calculated risk distribution. For the more recent minimum infection rate estimates (Boshuizen et al. 2006), the calculated rates are somewhat higher than the reported rates. The mean of the minimum infection rate is at approximately the 22nd percentile of the calculated distribution, with the lower and upper confidence limits on the

Figure 4 | Comparison of the calculated clinical severity risk with the reported clinical infection risk and estimated mortality risk for the West Frisian outbreak worker group at less than 15 metres from the whirlpool spa. Bar ends = 2.5 and 97.5th percentiles, box ends = 25 and 75th percentiles, ○ = median.

Figure 5 | Comparison of the calculated clinical severity risk with the reported clinical infection risk and estimated mortality risk for the West Frisian outbreak worker group at greater than 15 metres from the whirlpool spa. Bar ends = 2.5 and 97.5th percentiles, box ends = 25 and 75th percentiles, ○ = median.
The presumption that the model developed from the guinea pig mortality data predicts clinical severity disease in humans appears not to hold with risks calculated using the exposure assumptions applied for this whirlpool spa outbreak evaluation. As shown in Table 4, the reported rates for clinical severity infection fall above the 99th percentile of the calculated risk distribution for both the closer and further worker groups. However, the reported clinical infection rates are within an order of magnitude of the calculated risks. If mortality had occurred (which did not happen in the worker groups) at the mortality rates generally reported from other outbreaks, (5% to 15% of the clinically severity cases) then the calculated risk provides an improved estimate. For the exhibitor group who worked at less than 15 metres from the spa, the presumed mortality range falls at the 94th to 99th percentile of the calculated risk distribution. For the group who worked at greater than 15 metres from the spa, presumed mortality rate falls at the 70th to 99th percentile of the calculated risk distribution.

**Natural hot spring spa outbreaks risk calculations**

Table 5 shows the model evaluation results for the hot spring spa outbreaks. Figures 4–7 compare the sub-clinical and clinical estimates to the reported clinical infection rates and the mortality rates for the Miyazaki and Shizuoka outbreaks. None of the reports on these outbreaks included data for sub-clinical infection rates, but we estimated the sub-clinical infection rates for the natural hot spring spa outbreaks by using the ratio of sub-clinical to clinical rates from the West Frisian outbreak (Boshuizen et al. 2000; Boshuizen et al. 2001).

**Miyazaki prefecture outbreak**

Table 5 summarizes the estimates for the exposure dose (Armstrong 2005; Armstrong & Haas 2007a) and estimated Legionnaires’ disease infection risks. Figures 4 and 5 compare the sub-clinical and clinical modeled estimates to the reported clinical infection rate, the reported mortality rate and an estimate of the sub-clinical infection rate. Table 6 summarizes where the reported risks fall as percentiles of the calculated risk distributions.

The calculated clinical severity risk distribution reasonably predicts the reported clinical infection rate. The reported rate falls at the 86th percentile of the calculated rate distribution. The reported mortality is less than an order of magnitude below the calculated risk for clinical severity infection. The calculated sub-clinical risk distribution is within an order of magnitude of the estimate rate, which we derived from the clinical severity rate since the investigation report did not include this parameter, but the confidence intervals for the estimated rate do not overlap the calculated risk distribution.

**Shizuoka prefecture outbreak**

Table 5 summarizes the estimates for the exposure dose and the estimated Legionnaires’ disease infection risks. Figures 6 and 7 show the estimated rates with the 95th percentile confidence intervals compared to the estimated sub-clinical rate as well as the reported clinical severity infection rate and the reported mortality rate. The calculated sub-clinical
infection risk is higher than the estimated sub-clinical risk but is within an order of magnitude. The estimates of clinical severity risk overlap the reported clinical infection rate at the 99th percentile, as summarized in Table 6. However, for this outbreak report, the calculated clinical severity rates, although within an order of magnitude, are somewhat higher than the reported mortality.

**Legionella QMRA model validation results**

What constitutes an adequate validation of a quantitative chemical or microbial assessment appears to be somewhat subjective and not uniformly defined. The risk assessment literature widely discusses the need for validation, but with at best diffuse guidance. The main regulatory guidelines and web sites on risk assessment checked (EPA 1999; European Chemicals Bureau 2003; Health Canada 2004) provided nothing substantial on quantitative evaluation of the overall results, but provide some additional details on validation of components of a risk assessment, including the initial animal toxicology models, or software used in computational predictions of exposures.

The *Legionella* QMRA results agree generally within less than an order of magnitude with the rates from three different outbreaks, and this seemingly fits well within the general realm of validation. In chemical risk assessment, it is widely recognized that validation of the results is often difficult and the optimal approach may be new data generation, which may not always be feasible. There appear to be no rigorous guidelines on statistical evaluation of a validation exercise, beyond that given by sensitivity and uncertainty analyses (Burmaster & Anderson 1994). Few chemical risk assessments have adequate human data for the exposures which led to the observed risks and this limits validation. Unequivocal risk assessment validation is difficult using animal model extrapolation, unless convincing direct human data are available. However, if conclusive human data are available, then the animal extrapolation is unnecessary. This difficulty in validating a risk assessment holds even for a relatively well-studied contaminant with extensive human epidemiologic data, such as benzene. For benzene, the underlying exposure assessment components of the risk assessments arguably remain in debate, with the Pliofilm benzene-exposed cohort the subject of at least 4 distinct exposure assessments (Rinsky et al. 1987; Paustenbach et al. 1992; Crump 1996; Williams & Paustenbach 2003). These exposure assessments on the same workers and factories then led to differing epidemiologic analyses and subsequently differing risk assessments (Paustenbach et al. 1992).
This on-going debate suggests that, despite decades of research, the quantitative benzene risk assessments remain less than definitively validated and accepted.

Quantitative chemical risk assessments often utilize uncertainty factors, such as a factor of up to 10 for sensitivity differences in interspecies extrapolation and another factor of up to 10 for residual scientific uncertainty (Lewis et al. 1990; Dourson et al. 1996; Renwick & Lazarus 1998; Meek et al. 2002). These factors seemingly indicate that the overall validity may be no closer than within a factor of 10 to 100 of the actual risk. If so, then QMRA validation criteria of “within an order of magnitude” may also be reasonable. Taking that liberty, the current Legionella QMRA results as summarized in Figures 2–7 and Table 7 appear to be valid. We review the limits and qualifications on this conclusion in the Discussion section. Development of guidance on the quantitative validation of risk assessments seems to be an area meriting additional research.

**DISCUSSION**

**Strengths of the findings**

The results indicate that the models used for this Legionnaires’ disease QMRA are generally adequate to predict the disease experience for the outbreaks evaluated. All predic-
The QMRA mortality predictions, modeled from the healthy young guinea pigs in the various experimental mortality experiments, yield predicted rates which align well with the human experience. The human mortality from Legionnaires’ disease occurs dominantly in the elderly or those with susceptibility factors. Considering that the guinea pigs in the mortality experiments received no medical intervention (no antibiotics, no ventilators, no other supportive care), this does not necessarily mean the healthy young guinea pigs are innately more susceptible than are healthy humans lacking significant risk factors. In addition, the guinea pigs in the mortality experiments received relatively large bolus doses in the experimental range. Most of those experiments had a low dose of no less than 200 CFU. The guinea pig infection data, which was at lower doses than the mortality experiments, also yields QMRA predictions which align well with the human experience for the sub-clinical infection rates as measured by serology. The people in that sub-clinical infection study were perhaps generally healthier and overall younger than typical outbreak mortality cases. This may further support the guinea pig model’s relative comparability to human experience.

The West Frisian Floral Show whirlpool spa outbreak reports (Boshuizen et al. 2001; den Boer et al. 2002) seem to have reliable information on the number of exhibitors exposed at various distances, good data on their typical hours at work during the outbreak period, and also good information on their seroconversion prevalence. Although the reports do not give data on the Legionella content in the whirlpool spa, this is otherwise a particularly informative set of data, and provides currently unique information for the Legionnaires’ disease QMRA model evaluation.

The selected reports on outbreaks at natural hot spring spas (Anonymous 2003; Yabuuchi & Agata 2004; Okada et al. 2005) provide data on the number of visitors, the number of clinical cases, the number of fatalities, and data on the Legionella content of the spa water. These reports

Table 6 | Summary of percentiles the reported natural hot spring outbreak rates are for the estimated rate distributions

<table>
<thead>
<tr>
<th>Mortality</th>
<th>Clinical severity infection</th>
<th>Subclinical severity infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miyazaki</td>
<td>&lt; 1 (70 to 95)</td>
<td>&lt; 1**</td>
</tr>
<tr>
<td>Shizuoka</td>
<td>&lt; 1 (&lt; 1 to 19)</td>
<td>99 (84 to 99)</td>
</tr>
</tbody>
</table>

These metrics are estimated as discussed in Section 5.3. ** The upper 95% confidence interval on the rate does encompass the calculated distribution. The Miyazaki calculated clinical severity risk is lower than the reported clinical severity rate, but higher than the reported mortality. The Shizuoka calculated clinical severity risk is lower than the reported rate, but higher than the reported mortality.

Table 7 | Summary of results of the evaluation of the calculated versus reported risks for the three outbreaks used for model validation

<table>
<thead>
<tr>
<th>Outbreak</th>
<th>Subclinical severity infection</th>
<th>Clinical severity infection</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Frisian</td>
<td>√</td>
<td>Reported &lt; 10 × higher</td>
<td>√*</td>
</tr>
<tr>
<td>West Frisian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miyazaki</td>
<td></td>
<td></td>
<td>&lt; 10 × lower*</td>
</tr>
<tr>
<td>Shizuoka</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Based on estimated rates since the original investigators did not report these data.

√ = the confidence intervals of the predicted risks overlap the corresponding reported rates.
thus provide some unique information for use in the Legionnaires’ disease QMRA model evaluation. The two reports gave significantly different infection rates, and different Legionella counts in the spa water. The different modeled risk estimates mirror the different reported rates.

**Limitations of the findings**

For the model validation, we based the QMRA model risk projections on estimated exposures, since none of the outbreak reports provided information on case exposure measurements. It is possible that the exposure estimates are in error, despite efforts to use reasonable, data driven input parameter assumptions. However, comparison of the estimated air concentrations for the outbreaks compared reasonably with air concentrations of Legionella reported for aerosol samples from contaminated water sources (Armstrong & Hass 2007a).

The reported clinical severity infection cases from the outbreak investigations are from a somewhat imprecise endpoint, and inclusion or exclusion as cases depends on a number of investigative decisions. These include the criteria for including probable cases based on symptoms and potential for exposure, rather than by clinical diagnostics such as urinary antigen testing. The inclusion criteria may then lead to inaccurate case ascertainment. This case definition issue may not be severe in the West Frisian outbreak, where the seroprevalence data serves as one key indicator of infection. The number of clinical cases in all of the outbreaks does depend on the inclusion criteria. The number of clinical cases should nevertheless (arguably) be accurate to within an order of magnitude.

Another factor affecting the adequacy of the reported risk rates is the denominator (number of persons exposed) may be inaccurately determined, particularly for the hot spring spa outbreaks, where the number is likely to be the number of visits to the spa, and may not reflect the number of exposed individuals. If so, the actual rates of morbidity and mortality may be somewhat higher than reported, if the denominator for those exposed was actually a smaller number of individuals. The number of exposed workers in the West Frisian outbreak is a matter of record and captures the actual number of exposed workers. Thus, the denominator in the West Frisian outbreak risk rates is likely adequately determined.

A wealth of literature exists (perhaps numbering in the hundreds of journal articles) which investigate virulence factors for Legionella species and strains. In the QMRA model development, we assumed that the exposure is to a strain with virulence similar to that of the strains in the animal data used for the dose-response modeling. At present, we do not see any feasible way of further evaluating this assumption. The guinea pig infection data is in reasonable agreement with lowest dose group findings from the multiple investigations (Berendt et al. 1980; Baskerville et al. 1981; Davis et al. 1983; Fitzgeorge et al. 1983; Meenhorst et al. 1985; Jepras et al. 1985; Breiman & Horwitz 1987; Twisk-Meijssen et al. 1987) which reported guinea pig mortality experimental data, primarily for exposure to L. pneumophila serogroup 1 (Lp SG1). Lp SG1 is the most prevalent Legionella species and serogroup reported for Legionnaires’ disease. It is not currently known how other strains and species quantitatively compare in virulence to the strains of Lp tied to clinical disease. Qualitatively some other strains and species seem to be of comparable virulence, but others are likely less potent in causing morbidity or mortality. The QMRA model, based on L. pneumophila SG1, thus has somewhat uncertain applicability to the whole genus Legionella. Quantitatively establishing the relative virulence of different Legionella species and strains requires additional research. This current QMRA model for LD does not consider the potential role of protozoa on modulating Legionella virulence or viability, via intracellular adaptation or via vesicle production (O’Brien & Bhopal 1993; Berk et al. 1998). Further research is needed to consider the impact of these and other factors such as Legionella viability in aerosol and the role of human susceptibility factors.

A QMRA model may not be a significant tool in developing control strategies for Legionnaires’ disease, nor was that our intention when we undertook the model development. For prevention, improved strategies to prevent high-level source contamination are the most relevant course of action. A QMRA model for Legionnaires’ disease may nevertheless lead to better understanding the risk from Legionella exposure, and thus help motivate development of improved control guidance.
GENERAL CONCLUSIONS

The findings from this QMRA model evaluation for Legionnaires’ disease appear to have several general implications for QMRA of airborne pathogens. Typically, in chemical risk assessment, interspecies extrapolation factors adjust the animal model results for the usually presumed greater sensitivity of the heterogeneous human population. Use of interspecies adjustment factors has been a topic of some conjecture for QMRA. However, given a comparable animal model, the Legionnaires’ disease QMRA model findings suggest, for this application, no need for interspecies extrapolation factors. We suggest caution, however, since a review of data shows that different animal species models demonstrate orders of magnitude differences in resistance to infections (Armstrong 2005; Armstrong & Haas 2007b). An inappropriate animal model choice could give quite misleading results. For QMRA, an understanding of the comparative immunology of the candidate animal species and humans appears to be crucial. This is particularly so for a naive (previously unexposed to the organism) host’s initial cell-mediated immune system response to a pathogen, where significant interspecies differences in alveolar macrophage bactericidal mechanisms and invading bacterial evasion mechanisms can lead to major differences in host organism susceptibility.

Chemical risk assessments often include dose-scaling metrics, to account for body mass driven differences or other scale-based differences in response to a toxicant. The use of dose scaling in QMRA has been the subject of some conjecture. Mechanistic and probability based analysis of bacterial deposition, uptake and replication in alveolar macrophages suggests that dose scaling may be inappropriate for an intracellular pathogen in QMRA (Armstrong 2005; Armstrong & Haas 2007b). Findings from this QMRA, which do not include dose scaling, appear to support those theoretical arguments. Thus, this Legionnaires’ disease QMRA indicates no need for dose scaling for pathogenic organisms which replicate in the host.

The findings of this QMRA suggest that given a virulent organism, the probability of disease ties to the probability of exposure. This implies that a relatively low-level exposure to a virulent bacterial strain, widely distributed among a large population, will yield a burden of disease. For Legionnaires’ disease, this implies that, although high-level sources such as contaminated cooling towers can be significant in outbreaks, widespread exposure to lower level sources of a virulent strain could cause elevated LD rates in communities. Thus, high prevalence exposures to lower level sources may deserve further consideration in outbreak investigations.

Limited research has been done on the community risks of LD over time in relationship to cooling tower or other potential source proximity, although findings (Bhopal & Fallon 1991) suggest this is a contributor to the burden of LD. Given our LD QMRA findings, further study of the geographic clustering of LD cases over time may give insights into the impact of lower-level sources than those associated with frank LD outbreaks, and then give an impetus for revisions on Legionella control guidance.

ACKNOWLEDGEMENTS

The work presented is part of Dr. Armstrong’s research undertaken toward fulfillment of requirements for a doctoral degree completed at Drexel University. He received partial support for underlying components of the work from the ExxonMobil Mutualized Strategic Program.

REFERENCES

Armstrong, T.W. 2005 A Quantitative Microbial Risk Assessment Model for Human Inhalation Exposure to Legionella Doctor of Philosophy, Drexel University, Philadelphia, USA. Available at:http://dspace.library.drexel.edu/handle/1860/615
Armstrong, T. W. & Haas, C. N. 2007a A quantitative microbial risk assessment model for Legionnaires disease: assessment of...


Blanchard, D. C. 1989 The size and height to which jet drops are ejected from bursting bubbles in seawater. J. Geophys. Res. 94, 10999–11002.


Fitzgeorge, R. B., Baskerville, A., Broster, M., Hambleton, P. & Dennis, P. J. 1985 Aerosol infection of animals with strains of Legionella pneumophila of different virulence: comparison with intraperitoneal and intranasal routes of infection. J. Hyg (Lond) 90, 81–89.


First received 2 January 2007; accepted in revised form 14 March 2007. Available online January 2008