Available online at www.sciencedirect.com
ScienceDirect
Journal of Microbiological Methods 68 (2007) 184–192
Journal of Microbiological Methods
www.elsevier.com/locate/jmicmeth

A laboratory hot tub model for disinfectant efficacy evaluation

Darla M. Goeres *, Linda R. Loetterle, Martin A. Hamilton

Center for Biofilm Engineering, 366 EPS Building, Montana State University, Bozeman, MT 59717-3980, USA

Received 23 May 2006; received in revised form 13 July 2006; accepted 14 July 2006
Available online 1 September 2006

Abstract

This paper describes a novel laboratory hot tub (LHT) apparatus and associated standard operating procedure (SOP) designed to reproduce the key biological, chemical, and engineering parameters associated with recreational and therapeutic hot tubs. Efficacy, as measured quantitatively by log reduction values, was determined against both biofilm and planktonic bacteria. When the LHT was run according to the SOP, with no antimicrobial treatment, a consistent level of bacterial contamination occurred. The means of $\log_{10}$ viable cell densities (±the repeatability standard deviation of log densities) were 7.2 (±0.31) for the bulk water (density in units of cfu ml^{-1}), 5.3 (±0.56) for the coupons (density in units of cfu cm^{-2}), and 6.6 (±0.50) for the filters (density in units of cfu cm^{-2}). When control and chlorine treated LHTs were run in parallel, the log reduction increased significantly with chlorine concentration for samples of planktonic bacteria in the bulk water ($p=0.016$), biofilm bacteria on the coupons ($p=0.09$) and biofilm bacteria on the filter ($p=0.005$), indicating that the method was sensitive to chlorine concentration. The method also displayed sensitivity by differentiating between chlorine and bromine treatments; in every case, chlorine produced a greater log reduction than did the same concentration of bromine. The model and SOP were shown to be rugged with respect to slight changes in fluid mixing intensity, water chemistry (saturation index), inoculum size, and organic loading. The LHT and associated SOP provide a reliable second tier in a three-tiered testing process, in which the first tier is a suspension test and the final tier is a field test.

Keywords: Biofilm; Laboratory model; Log reduction; Repeatable test; Rugged method; Spa

1. Introduction

Hot tubs, also known as spas or spa pools, are approximately 1400 l pools of hot water used for recreational and therapeutic soaking. Hot tub standards recommend draining, cleaning, and refilling public hot tubs at a water replacement interval (days) equal to (1/3)(spa volume, US gallons)/(max no. of daily uses) or when the total dissolved solids (TDS) in the water exceeds the source water TDS by 1500 ppm (ANSI/NSPI-2, 1999). In the interim, the water quality is maintained by the addition of chemicals and recycling the water through a skimmer and filter. Recommended ranges for oxidizing disinfectants, stabilizers, pH, alkalinity, calcium hardness, and TDS are found in the American National Standards for Public Spas (ANSI/NSPI-2, 1999), American National Standards for Portable Spas (ANSI/NSPI-6, 1999) and American National Standards for Permanently Installed Residential Spas (ANSI/NSPI-3, 1999). Guidelines for the skimmer and filter design are found in Circulation System Components and Related Materials for Swimming Pools, Spas/Hot tubs (ANSI/NSF 50, 2001). Compliance rules for bacterial contamination in hot tubs are set by local health departments.

Maintaining a balanced water chemistry in a hot tub is challenging for three reasons. Bathers introduce an organic load on the water chemistry, thereby neutralizing the disinfectant within the water (Judd and Black, 2000; Kim et al., 2002). The water in a hot tub may reach a temperature of 40 °C (ANSI/NSPI-2, 1999), which affects the kinetics. Lastly, hot tubs are outfitted with hydrojets and air blowers to enhance the bathing experience; these devices thoroughly mix and oxygenate the water.

Hot tub users may be exposed to waterborne pathogens and are at risk of illness, including respiratory infections (Kahana et al., 1997; Mangione et al., 2001; Rickman et al., 2002), Pontiac Fever (Fallon and Rowbotham, 1990; Fields et al., 2001; Spitalny et al., 1984a), gastrointestinal disease (Holmes et al., 1989), urinary tract infections (Salmen et al., 1983), or ear infections (Havelaar et al., 1983), eye infections (Insler and Gore, 1986; Samples et al., 1984) and skin infections (Gustafson et al., 1983; McCausland

* Corresponding author. Tel.: +1 406 994 2440; fax: +1 406 994 6098.
E-mail address: darla_g@erc.montana.edu (D.M. Goeres).

0167-7012/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.mimet.2006.07.010

and Cox, 1975; Spitalny et al., 1984b). Folliculitis that occurs as a result of hot tub use, which accounts for the majority of hot tub illness related reports, are most often linked to *Pseudomonas aeruginosa*. More recent reports indicate the importance of nontuberculous *Mycobacteria* (Mangione et al., 2001) infections in hot tubs. The presence of *Legionella* in hot tubs has been observed and associated with Pontiac Fever or Legionnaires's Disease (Thomas et al., 1993).

The U.S. Centers for Disease Control and Prevention tracks recreational water disease outbreaks and reports their findings in the "Surveillance for Waterborne-Disease Outbreaks, United States" (http://www.cdc.gov/mmwr/). Even though diseases associated with hot tub use are well-documented in the literature, many incidences acquired in private hot tubs are probably never reported, especially in the case of folliculitis, which is self-limiting and often does not require medical attention.

The risk of infection and disease can be reduced by proper application of effective hot tub disinfectants. To provide assurance that a disinfectant is effective, federal regulatory authorities in the US require that hot tub disinfectants pass both the laboratory suspension test against each of *Escherichia coli* and *Streptococcus faecalis* (AOAC, 1990) and a field test (EPA, 1979). The AOAC method was developed to test the "germicidal activity" of swimming pool disinfectants (Ortenzio and Clark, 1964). When hot tubs became a part of the recreational water market, the swimming pool standards were applied even though hot tubs are different from swimming pools in many important ways. Hot tubs operate at higher temperatures, are more aerated due to the hydrojets and air blowers, have a greater surface area to volume ratio, serve more bathers per water volume under typical use, and more frequently recycle the water through filters. Based upon these differences, there has been a call for a new set of standards designed specifically for hot tubs (Crandall and MacKenzie, 1984; Kush and Hoadley, 1980). It is important to note that the currently accepted methods do not monitor or measure the extent of biofilm growth.

Biofilm bacteria may well be a significant source of microbial contamination in hot tubs (Price and Ahearn, 1988). Biofilms often harbor pathogens (Murga et al., 2001; Primm et al., 2004). Bacteria in biofilms are more tolerant to treatment than planktonic bacteria (Costerton et al., 1999; Highsmith and Favero, 1985; Lewis, 2001). Bacteria in detached biofilm particles can be more pathogenic than individual planktonic bacteria; e.g., in causing a characteristic pulmonary injury in guinea pigs, 10 to 100 times more suspended cells of *Legionella pneumophila* were required than if the cells were contained in artificial microcolonies (Wright, 2000). For these reasons, there is a pressing need for a laboratory hot tub disinfectant test against biofilm bacteria as well as planktonic bacteria (Favero, 1984; Hamilton, 2002).

This paper presents a laboratory hot tub (LHT) system and an associated standard operating procedure (SOP) that together are suitable for assessing the efficacy of hot tub disinfectants against biofilm and planktonic bacteria. A series of experiments was conducted to evaluate the practicality, repeatability, sensitivity, and ruggedness of the LHT and SOP.

2. Materials and methods

2.1. Laboratory hot tub (LHT)

The LHT was designed to represent the engineering, water chemistry and microbial components of a 1420 l hot tub built to accommodate four people. In a typical hot tub, the water passes through a skimmer which removes the larger particles, a filter which removes smaller particles, and a heat exchanger to maintain the temperature within a specified range. Then the water recycles back into the tub. The hot tub water must make this cycle every 30 min. ANSI/NSF recommends that surface-type filters for use in public hot tubs should receive a maximum loading of 1.5 ml min^{-1} cm^{-2} (ANSI/NSF 50, 2001).

The tub component of the LHT consisted of a 1-l beaker, called a reactor, that possessed no spouts or pour lips. A 1-in. stir bar was placed in the bottom of the reactor (Fig. 1a) for gentle mixing. Two reactors were placed on a submersible stir plate with digital rpm control in a constant temperature water bath maintained at 38 °C (Fig. 1d). One reactor was randomly selected to be the untreated control and the other reactor received a treatment. The efficacy of the treatment was based on a comparison of the microbial accumulation (planktonic and biofilm) in the control reactor with the ostensibly lower accumulation in the treated reactor. By running the control and treated reactors simultaneously the comparison was not influenced by inherent, unknown, or uncontrolled factors that change from day to day and potentially affect microbial accumulation.

Each reactor was capped with a #15 rubber stopper that contained four barbed ports and six holes sized to fit #3 rubber stoppers (Fig. 1a). Each #3 stopper contained a Teflon™ rod fitted with a 1 cm^3 polycarbonate coupon for biofilm sampling.

Ports 1 and 2 were used for air exchange. A piece of tubing with an air sparging stone was connected to the bottom side of port 1 and an 0.2 μm air filter followed by an air pump was connected to the top side of port 1. The air pump was the same style that is used in household fish tanks and was included to simulate the hydrojets. A 0.2 μm air filter was connected to port 2 to provide an escape route for the sparged air.

Ports 3 and 4 were the effluent and influent for the filter recycle loop. A piece of tubing that extended to the bottom of the reactor was connected to the bottom of port 3. Three Y connectors were attached to the top of port 3 in such a way that four individual lines resulted (Fig. 1b). A minimum amount of tubing was used for this configuration. Each of the four lines were then fed through a pump head. All four pump heads were driven from one pump. An on-line filter (Pall Gelman Laboratory; 25 mm in-line filter holder, Delrin) was attached to the end of the tubing that came out of the pump head. Each on-line filter contained a piece of commercial hot tub filter material (Unicel C-6430) cut into a circle with a 2.5 cm diameter where the effective filter diameter was 2.17 cm (Fig. 1c). The total filter surface area was 14.8 cm^2. Only new commercial hot tub filter material that contained no chemical or biological residue was used. Following the filters, the Y connectors were reversed and hooked to the top of port 4 and a piece of tubing located just below the liquid/air interface was connected to the bottom of port 4. To achieve a filter loading rate of 1.02 ml

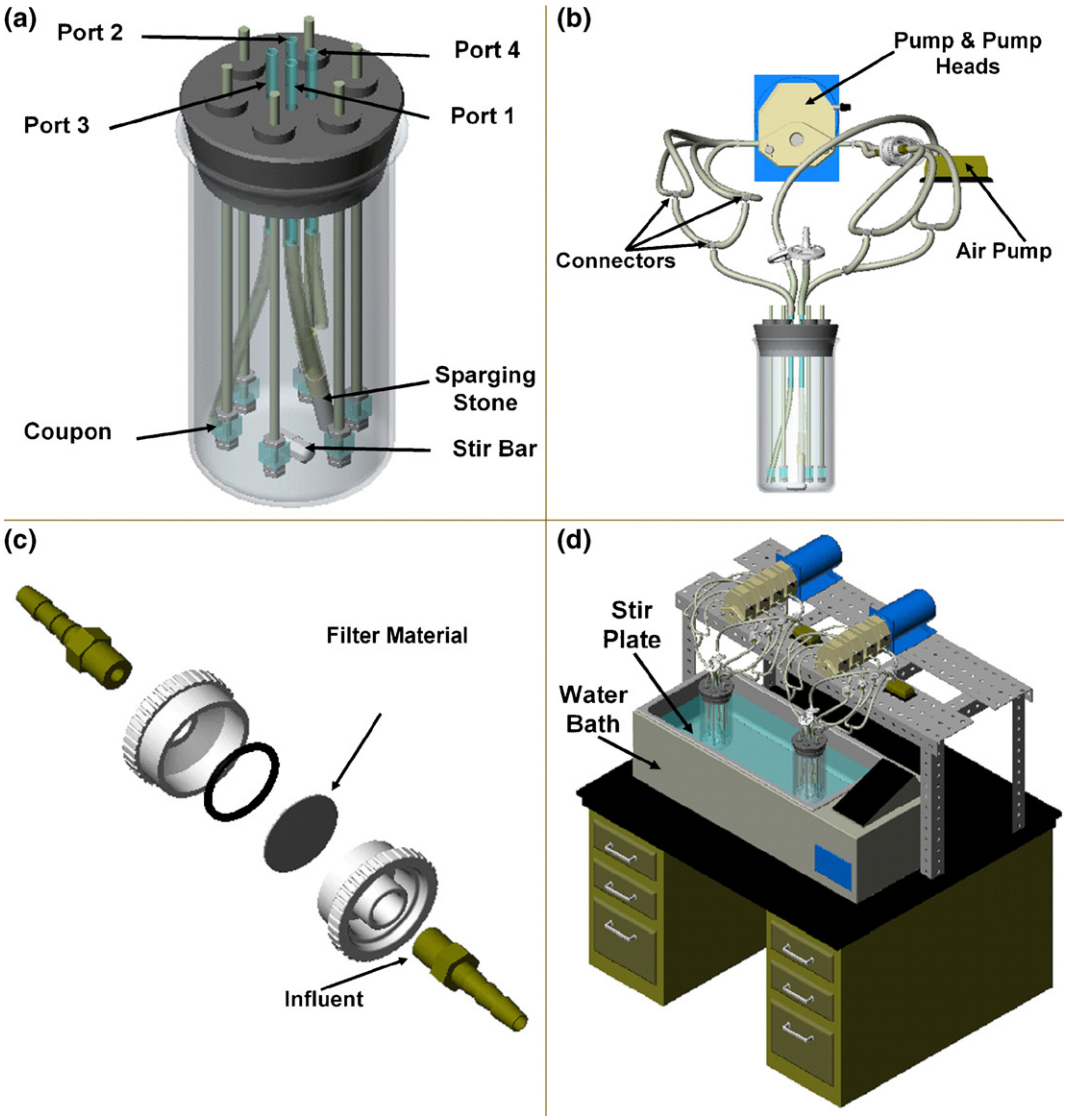

Fig. 1. Schematics of the various components that comprise the laboratory hot tub system. Panel (a) illustrates the ports, coupons, sparging stone and stir bar. Panel (b) demonstrates how the laboratory reactor is plumbed. Panel (c) is a blown-up schematic of one filter. Panel (d) is an illustration of the complete system as it rests in a water bath on a stir plate.

min^{-1} cm^{-2} and 30 min residence time required an overall flow rate of 15.1 ml min^{-1} through the four filters and a reactor volume of 450 ml. Collectively, both reactors, tubing, filters, water bath and pumps were labeled "the LHT system."

2.2. Standard operating procedure (SOP)—preparation

Before the start of an experiment the reactors and associated tubing were cleaned. The system was assembled, minus the filters, a 1% bleach solution was added and the pumps were turned on for 30 min. After the bleach soak, 10 ml of a 78,750 mg l^{-1} sodium thiosulfate stock solution was added and the pumps turned on for another 30 min. The liquid was drained and all pieces thoroughly rinsed with tap water. A scrub brush and soap were used to remove any residual debris on the reactor pieces and to wash the filter holders. This process was followed by a thorough rinsing. If any reactor piece was visibly fouled, it was sonicated for 10–15 min for extra cleaning or replaced. Each piece was then allowed to air dry.

Twenty-four hours before the reactors were inoculated, they were filled with 450 ml of nano pure water and autoclaved. The filters were sterilized separately and the entire system assembled once the pieces cooled. The assembled system was placed in the 38 °C water bath on top of the stir plate and the tubing was fed through the pump heads and the air pump was connected to the end of the filter attached to port 1. The stir plate was adjusted to rotate at 130 rpm.

The water chemistry in the laboratory hot tub was adjusted to fall within the ANSI/NSPI recommended guidelines for a public hot tub (ANSI/NSPI-2, 1999) which achieved a near neutral saturation index (SI). The calcium hardness and alkalinity and were manipulated to equal 200 mg $CaCO_3$ l^{-1} and 120 mg $CaCO_3$ l^{-1}, respectively, in the bulk fluid with the addition of 1.35 ml of an autoclaved 31.74 g $MgCl_2$ l^{-1} and 73.99 g $CaCl_2$ l^{-1} solution and

a 1.40 ml addition of a filter sterilized 56.03 g $NaHCO_3$ l^{-1} solution (AOAC, 1995). Using aseptic technique, the pH was adjusted to equal 7.2. At this point, one reactor was randomly designated as a treatment reactor and the other as the control reactor.

2.3. Standard operating procedure—disinfection

Chlorine, added as sodium hypochlorite, was used as the model disinfectant during SOP development, repeatability testing and ruggedness testing. For the sensitivity analysis both chlorine and bromine, added as a powder that contained active ingredients of sodium dichloro-*s*-triazinetrione and sodium bromide, were utilized.

Twenty-four hours prior to inoculating the reactors, disinfectant was added to the treated reactor to meet the demand the system itself placed on the disinfectant. For chlorine, the demand was generally met with 5–10 mg free available chlorine (FAC) per liter. The demand was successfully met when a measurable disinfectant residual was found. At the end of 24 h, the disinfectant concentration in the treated reactor was measured, and if necessary, the concentration was adjusted to the target value.

2.4. Standard operating procedure—inoculation and organics

At time zero, each reactor was inoculated with a mixture that contained 0.1 ml of a 10^8 cfu ml^{-1} suspended culture of *P. aeruginosa*, originally isolated from a hot tub, and 1.0 ml of a sterile solution that contained 9.0 g tryptic soy broth (TSB) l^{-1} and 3.6 g urea l^{-1}. This concentration of *P. aeruginosa*, undefined carbon and urea is consistent with the ANSI/NSPI 50 test method (Annex H of ANSI/NSF 50, 2001). Although *P. aeruginosa* is not necessarily a component of normal human flora, it is often associated with hot tubs and hot tub diseases (Highsmith and Favero, 1985; Highsmith and McNamara, 1988). Immediately, the air pump was started and both reactors sparged for 20 min.

2.5. Standard operating procedure—sampling and analysis

At the end of 20 min, the air pump was turned off and the disinfectant concentration was measured in the treated reactor. At 24 h post inoculation, a second disinfectant concentration was measured in the treated reactor; also, a bulk water, two coupons and four filter samples were collected from both the treated and control reactors and analyzed for viable bacteria. Bacterial samples were collected after 24 h to allow time for injured cells to recover and biofilm to form. The hardness, alkalinity and pH were then measured.

Alkalinity was measured with a Hach test kit (Cat. No. 24443-01) based upon the sulfuric acid titration method (APHA, 1995). pH measurements were read on a calibrated pH meter. Free available chlorine (FAC) was measured according to the DPD method (APHA, 1995) using a standard curve prepared with Hach free chlorine packets (Cat. No. 14077-99) and potassium permanganate. Bromine was measured according to the DPD method (APHA, 1995) using a standard curve prepared with Hach total chlorine packets (Cat. No. 14076-99). Calcium hardness was measured with a Hach test kit (Cat. No. 1457-01) based upon the EDTA titration method (APHA, 1995).

To measure the viable cell counts in the LHT bulk water, samples were serially diluted in sterile buffered water. The first dilution tube for both the treated and control samples contained sodium thiosulfate to neutralize the chlorine or bromine. The samples were plated on R2A agar, incubated for 18–24 h at 37±2 °C and enumerated.

Coupon biofilm samples were collected by lifting a randomly selected rod and coupon through the stopper lid on the beaker. The coupon was gently dipped in sterile dilution water to remove any unattached cells. The side of the coupon that faced towards the center of the beaker was scraped into a dilution tube containing 9 ml sterile buffered water that contained sodium thiosulfate and then was rinsed with 1 ml of sterile buffered water. The suspension of removed biofilm bacteria was sonicated for 2 min at 50/60 Hertz. Samples were then serially diluted, plated on R2A agar, incubated for 18–24 h at 37±2 °C and enumerated.

Filter biofilm samples were collected by placing the filter into 10 ml sterile buffered water containing sodium thiosulfate. The sample was sonicated for 30 s and vortexed for 30 s for two cycles. Finally, the sample was serially diluted, plated on R2A agar, incubated for 18–24 h at 37±2 °C and enumerated.

2.6. Statistical analysis

Each viable cell count was expressed as a density, where the units were cfu ml^{-1} for a bulk water sample, cfu cm^{-2} for a coupon, and cfu cm^{-2} for a filter. For a filter, the surface area was measured at the macroscale, ignoring the fact that a large microscale surface area was available on the individual fibers (Fig. 2). All statistical analyses were conducted on $\log_{10}$ transformed

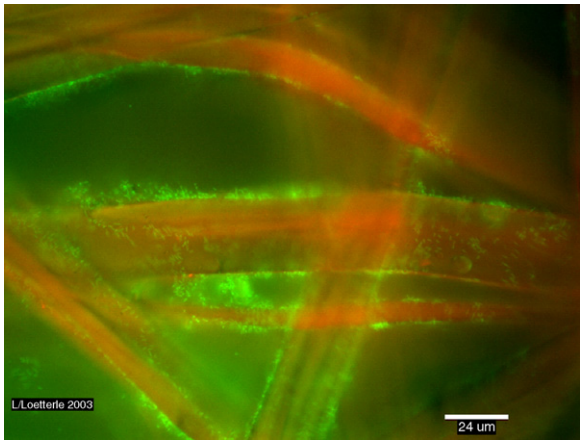

Fig. 2. An epiflorescent micrograph of biofilm on the fibers of hot tub filter material. The filter was sampled from a control LHT system at the conclusion of a SOP run. The bacteria in the biofilm were stained with *Bac*Light Live Dead (Molecular Probes, Eugene, Oregon). The live bacteria display in green and the hot tub fibers in orange. The fibers that make up a hot tub filter provide a large surface area for biofilm growth. The filter may well provide a reservoir of viable bacteria that can contaminate an otherwise thoroughly cleaned and disinfected hot tub.

Table 1
Ruggedness test parameter settings

Settings	Mixing (rpm)	Water chemistry	Inoculum ($\log_{10}$ cfu ml^{-1})	Organic contribution (mg l^{-1})
Standard operating procedure (SOP)	130 1 in. stir bar	SI −0.06	7	20 TSB 8 urea
Low	200 1 in. stir bar	SI −0.70	6	10 TSB 4 urea
High	200 1/2 in. stir bar	SI +0.70	8	40 TSB 16 urea

densities (log densities). The control log densities were analyzed to determine the ability of the LHT system to produce consistent bacterial contamination when there was no application of a disinfectant. The log reduction, which is the mean log density for the control LHT minus the mean log density for the parallel treated LHT, was the disinfectant efficacy measure (Zelver et al., 2001). The repeatability and ruggedness of the LHT system was evaluated for both the control log densities and the log reduction values. The sensitivity of the log reduction values was also assessed.

A method is said to be "repeatable" if the quantitative response exhibits a small standard deviation when the method is independently repeated (IUPAC, 1997). The repeatability experiments were conducted by three technicians who worked in two separate laboratory rooms at the Center for Biofilm Engineering, Bozeman, MT. Technician 1 completed six experiments, technician 2 completed 14 experiments and technician 3 completed 16 experiments for a total of 36 experiments. Repeatability was evaluated by submitting the quantitative responses to an analysis of variance (ANOVA) where laboratory by technician combinations, experiments, and samples were sequentially nested random effects factors.

A method is considered "rugged" if slight changes to the SOP induce only a small effect on the quantitative response (Goeres et al., 2005). By emulating the small departures from the SOP that may well occur when the method is implemented in different laboratories, ruggedness experiments can produce an estimate of interlaboratory variability. In this study, the SOP was altered to create two settings of the operational factors, a low setting chosen to discourage microbial growth and a high setting chosen to encourage microbial growth (Table 1). The LHT was run 5 or 6 times at each setting using a free available chlorine concentration of 3 mg l^{-1}. An ANOVA was used to calculate the standard error of each quantitative response at the low and high settings. For the ANOVA, the high or low setting was a fixed effects factor and experiment and sample were sequentially nested random effects factors. The difference between means, high setting minus low setting, was evaluated by conducting a two-sample *t*-test and finding the associated upper one-sided confidence limit.

A method is considered "sensitive" if it is able to produce a concentration by efficacy positive trend (IUPAC, 1997). The LHT was run against three concentrations of free chlorine (1, 3 and 5 mg l^{-1}) and two concentrations of bromine (1 and 3 mg l^{-1}). A trend test based on the slope of a least squares regression

Table 2
Statistical summary of the log densities in the control reactors when operated according to the SOP

	Mean	Repeatability standard deviation	Percentage contribution to the total variance		
			Laboratory and technician	Experiment	Sample
Bulk water $\log_{10}$ cfu ml^{-1}	7.2	0.31	40%	60%	NC[a]
Coupons $\log_{10}$ cfu cm^{-2}	5.3	0.56	45%	34%	21%
Filters $\log_{10}$ cfu cm^{-2}	6.6	0.50	28%	56%	16%

[a] Not calculable because only one bulk water sample was collected for each experiment.

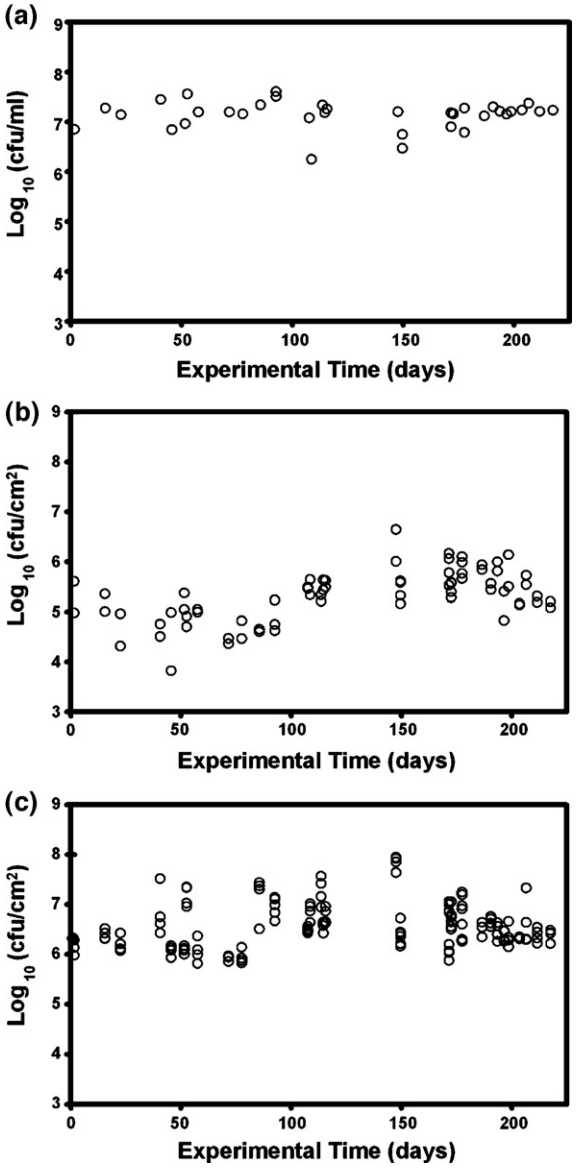

Fig. 3. Time series plots of control bacterial counts illustrating the time trend associated with the method. Data are pooled for all technicians and both laboratories. Panel (a) contains the bulk water data, panel (b) the coupon data and panel (c) the filter data. Points aligned vertically are either samples collected from different experiments completed on the same day and/or multiple samples collected from the same experiment.

Table 3
Statistical summary of the log reduction values associated with each chlorine treatment when the LHT system was operated according to the SOP

	Chlorine concentration (mg l^{-1})	Mean	Repeatability standard deviation	Percentage contribution to the total variance	
				Laboratory and technician	Experiment
Bulk water	1	5.1	2.56	89%	11%
$\log_{10}$ cfu ml^{-1}	3	6.5	0.67	10%	90%
	5	7.3	0.35	51%	49%
Coupons	1	3.9	2.21	47%	53%
$\log_{10}$ cfu cm^{-2}	3	4.8	0.55	44%	55%
	5	5.0	0.48	0	100%
Filters	1	4.2	2.93	86%	14%
$\log_{10}$ cfu cm^{-2}	3	5.5	1.24	61%	38%
	5	6.5	0.21	38%	62%

The numbers of independent experiments were $n=5$, 7, and 4 for chlorine (FAC) concentrations of 1, 3, and 5 mg l^{-1}, respectively.

line was used to evaluate the sensitivity of the log reduction values to increasing chlorine concentration. An analysis of covariance was used, where disinfectant concentration was the regression predictor and laboratory, technician and experiment were nested random effects factors.

An alternative definition of sensitivity is the ability to produce the correct ordering of log reduction values when testing two disinfectants known to have different efficacies. A two-factor main effects ANOVA model, where disinfectant and laboratory were fixed effects factors, was used to compare the log reduction values associated with chlorine versus bromine.

All statistical analyses were done using the computer package Minitab, Release 13 (Minitab, Inc., State College, PA). A 95% confidence level was used when calculating an upper, one-sided limit and a 90% confidence level was used for a two-sided interval.

3. Results

Table 2 shows the means, repeatability standard deviations, and variance components for the control log densities. The repeatability standard deviations ranged from 0.31 for the suspended bacteria in the bulk water to 0.56 for biofilm bacteria on coupons. No consistent pattern was found for the percentage contributions from the three sources—among combinations of laboratories and technicians, among experiments, and among samples within an experiment (Table 2).

Fig. 3 displays the variability of control log density values observed during an 8-month period. There was no time trend for the bulk water ($p=0.9$) and filters ($p=0.2$). The coupon data, Fig. 3b, had a statistically significant trend ($p<0.001$), but it amounted to an increase of only 1 cfu cm^{-2} day^{-1} which is not of practical importance.

Table 3 displays the means, repeatability standard deviations, and variance components for the log reductions. The log reduction repeatability standard deviation was always greatest at the lowest treatment concentration and lowest for the highest treatment concentration. For the log reduction data, the variability was partitioned into two sources—among combinations of laboratory and technician and among experiments. No consistent pattern was discerned for the percentage contributions from the two sources.

Table 4 presents the ruggedness results for the control data. The large p-values indicate that there were no statistically significant differences between the high and low settings. One can be 95% confident that the difference in true mean control log densities, high setting minus low setting, was less than 0.65 for bulk water samples, less than 0.88 for coupon samples, and less than 0.97 for filter samples (Table 4).

Table 5 presents the ruggedness results for log reductions. The difference in mean log reductions, high setting minus low setting, could conceivably be positive or negative. The confidence intervals in Table 5 show that the data do not discredit a difference between log reductions as large as −2.9 for coupons.

Table 4
Statistical summary of the bacterial log densities in the control reactors for each ruggedness test setting

Sample	Mean for growth setting		SEM	2-tailed p-value	95% upper, one-sided confidence limit
	Low	High			
Bulk water $\log_{10}$ cfu ml^{-1}	7.1	7.4	0.12	0.1	0.65
Coupons $\log_{10}$ cfu cm^{-2}	5.5	5.9	0.19	0.2	0.88
Filters $\log_{10}$ cfu cm^{-2}	6.7	7.0	0.24	0.4	0.97

The standard error of the mean (SEM) applies to both means in the row. The p-value and upper confidence limit pertain to the true log density mean difference, high setting minus low setting.

Table 5
Statistical summary of log reductions due to a chlorine treatment of 3 mg l^{-1} as observed for each ruggedness test setting

Sample	Mean for growth setting		SEM	2-tailed p-value	90% confidence interval
	Low	High			
Bulk water $\log_{10}$ cfu ml^{-1}	7.1	6.3	0.48	0.3	−2.1, 0.4
Coupons $\log_{10}$ cfu cm^{-2}	5.3	4.2	0.68	0.3	−2.9, 0.7
Filters $\log_{10}$ cfu cm^{-2}	6.4	6.4	0.43	0.96	−1.1, 1.1

The standard error of the mean (SEM) applies to both means in the row. The p-value and confidence interval pertain to the true mean log reduction difference, high setting minus low setting.

Table 6
Comparison of log reduction values calculated for chlorine and bromine, each tested at two concentrations

	Disinfectant concentration (mg l^{-1})	Mean for chlorine	Mean for bromine	2-tailed p-value
Bulk water	1	5.5	3.7	0.004
$\log_{10}$ cfu ml^{-1}	3	6.4	5.0	0.08
Coupons	1	4.4	2.6	0.06
$\log_{10}$ cfu cm^{-2}	3	4.5	3.5	0.1
Filters	1	4.8	2.3	0.007
$\log_{10}$ cfu cm^{-2}	3	5.3	4.3	0.08

The p-values pertain to the true difference in mean log reductions, chlorine minus bromine.

For the sensitivity analysis, regression trend tests showed that the log reduction increased with chlorine concentration for the bulk water ($p=0.016$), coupons ($p=0.09$) and filter ($p=0.005$). These results indicate that the method can detect the effect of increased chlorine concentration.

Table 6 compares the mean log reductions resulting from a chlorine treatment and a bromine treatment, both tested at 1 and 3 mg l^{-1}. The calculated p-values pertain to the null hypothesis that there was no difference between the chlorine and bromine treatments. In all cases tested, the chlorine resulted in the greater log reduction, indicating greater efficacy, although not all cases achieved statistical significance. For both the chlorine and bromine, the log reduction was always higher for 3 mg l^{-1} than for 1 mg l^{-1}.

4. Discussion

Empirical evaluations of the LHT and SOP demonstrated that the method was repeatable, rugged, and sensitive. The control log densities exhibited small repeatability standard deviations indicating that the LHT system creates relatively consistent planktonic and biofilm bacterial contamination in the absence of disinfection. No single factor consistently provided the predominate source of variability implying there is no obvious focal point for improving the system.

The statistically significant time trend associated with the coupon control log densities may well be attributable to technician experience. The coupons required the most difficult manipulations, especially for removing and disaggregating the biofilm. As the technicians became more proficient, it is plausible that the number of recovered biofilm bacteria correspondingly increased. However, because the rate of increase was of no practical importance, neither the method nor technician training requires modification.

The repeatability standard deviations for log reduction values were consistent with other disinfection tests (Tilt and Hamilton, 1999), except at the marginal chlorine treatment of 1 mg l^{-1} where the repeatability standard deviations were 2.6, 2.2, and 2.9 for bulk water, coupon, and filter samples, respectively. A high repeatability standard deviation for marginal treatments is consistent with repeatability results for standard disinfection tests (Tilt and Hamilton, 1999). The rationale for this observation is that an effective treatment consistently kills all bacteria so that the variability in the log reduction is mostly due to the control data. An ineffective treatment is essentially the same as the control so the standard deviation is similar to that for control data. Ineffective and very effective treatments represent points on opposite ends of a concentration by log reduction curve, where the slope of the curve is not changing dramatically. In contrast, marginal treatments correspond to points on the steepest part of the curve where influential, but unidentified and uncontrolled, factors cause increased variability.

The results imply that more replicate tests are required for a marginal treatment in order to average out the greater variability and to arrive at a precise estimate of efficacy. For example, the results of this investigation indicate that 16 tests of a marginal chlorine treatment such as 1 mg l^{-1} would be required to produce as precise a log reduction for biofilm bacteria on coupons as would a single test of an effective chlorine treatment such as 3–5 mg l^{-1}.

In spite of the low repeatability for marginal treatments, the regression trend test demonstrated that the laboratory hot tub method was able to distinguish between treatments of high and low efficacy. The sensitivity of the LHT suggests that it would be possible to craft a meaningful pass/fail performance standard for the LHT log reduction result.

Chlorine was chosen as the model disinfectant because of its historical prominence in the recreational water industry (AOAC, 1990; EPA, 1979; Brown and McLean, 1966). Bromine was chosen as the comparison disinfectant because of its accepted use as a disinfectant in recreational water. The method was able to show that chlorine had a greater efficacy than an equivalent concentration of bromine. This result is consistent with what other researchers have found (Koski et al., 1966; Shaw, 1984).

The engineering design of the LHT could not emulate every relevant characteristic of a real hot tub. In particular, the LHT had about an order of magnitude greater surface area to volume ratio than do real hot tubs. It was impossible to make the surface area to volume ratio for the laboratory model the same ratio as for field systems if the other engineering specifications were to be maintained. In biofilm reactor design, the surface area to volume ratio is important because biofilm forms on every available surface in the reactor, not just on the surfaces intended for sampling. The greater the surface area, the more biofilm is present to exert a demand on the disinfectant. The total amount of disinfectant (mg) present in the bulk liquid is calculated by multiplying the disinfectant concentration (mg l^{-1}) by the total liquid volume (l) in the reactor. Guidelines recommend an acceptable range for disinfectant concentration in hot tubs (ANSI/NSPI-2, 1999; ANSI/NSPI-3, 1999; ANSI/NSPI-6, 1999). For a set disinfectant concentration, less bulk liquid volume means less total disinfectant present to kill viable biofilm bacteria. For this reason, the treated beaker was conditioned for 24 h with an excess of disinfectant so that the system's additional demand, a result of the greater surface area to volume ratio, was met before the experiment began. In theory, the greater surface area to volume ratio in the laboratory reactor implies that it should provide a more conservative estimate of a disinfectant's efficacy. In practice, the surface area to volume ratio is only one parameter of many that impact a disinfectant's efficacy.

In the LHT, the biofilm coupons were manufactured from polycarbonate, whereas many commercial hot tubs are manufactured from fiber glass. This choice of coupon material for purposes of evaluating the LHT method is judged to be unimportant because previous research has shown that coupon construction material was not a critical design criteria for the laboratory evaluation of recreational water disinfectants (Goeres et al., 2004). The method can be easily modified to accommodate alternative coupon materials.

Information in this paper focused exclusively on an empirical evaluation of the LHT and SOP as a laboratory system potentially suitable for testing the efficacy of hot tub disinfectants. However, the method is also useful for conducting research

on hot tub control and operation strategies. The authors have conducted experiments with the LHT to investigate a variety of issues, including the rate at which regrowth occurs in a cleaned and disinfected hot tub due to biofilm on the filters, the relationship between biofilm accumulation and planktonic viable cell counts, the effectiveness of treatments administered to kill bacteria and/or kill and remove biofilm from a fouled system, and the efficacy of chlorine against *E. coli* and *Mycobacterium fortuitum* (data not reported). In the future, the LHT will be used to measure disinfection by-products that are generated when organics and chlorine mix, with and without aeration.

5. Conclusions

It is important to include biofilm contamination as a response variable in both laboratory and field hot tub tests. Biofilms can and do form in hot tubs, and once established are extremely difficult to remove. To ensure quality and safe water, future chemistries must address the control, kill, and removal of biofilms from hot tubs. The LHT was designed to mimic actual hot tub design and use conditions and it should provide an indication of how well a treatment is going to work in the field. We suggest that the laboratory hot tub model be used as a second tier in a three-tiered process, in which the first tier is the AOAC suspension test and the final tier is the currently required field test, modified to include biofilm measurements.

Acknowledgments

This work was partially supported by the US EPA Antimicrobials Division under Contract 68-W-02-050 with Montana State University-Bozeman. This document does not necessarily reflect the opinions of the EPA. The authors thank the following people for their contributions and for capably performing the many experiments conducted during this project: Kelli Buckingham-Meyer, Joanna Heersink, Ivy Able, Marion Clark, Liv Hyatt, Greg Mitchell, Melanie Nowlan, Liz Sandvik, Fred Siebert and Kristin Susens. The authors thank Patrick Norris for his drawings of the apparatus.

References

ANSI/NSPI-2, 1999. American National Standard for Public Spas. National Spa & Pool Institute, Alexandria, VA.

ANSI/NSPI-3, 1999. American National Standard for Permanently Installed Residential Spas. National Spa & Pool Institute, Alexandria, VA.

ANSI/NSPI-6, 1999. American National Standard for Portable Spas. National Spa & Pool Institute, Alexandria, VA.

ANSI/NSF 50, 2001. Circulation System Components and Related Materials for Swimming Pools, Spas/Hot Tubs. NSF International, Ann Arbor, MI.

AOAC, 1990. AOAC Official Method 985.13 Disinfectants (water) for swimming pools. AOAC Official Methods of Analysis. AOAC International, Gaithersburg, MD.

AOAC, 1995. AOAC Official Method 960.06 Germicidal and detergent sanitizing action of disinfectants. AOAC Official Methods of Analysis. Section E, Synthetic Hard Water. AOAC International, Gaithersburg, MD.

APHA, 1995. Standard methods for the examination of water and waste waters, 19th ed. Eaton, A.D., Clesceri, L.S., Greenberg, A.E. (Eds.), American Public Heath Association, American Water Works Association, Water Environment Federation, Washington, DC.

Brown, J.R., McLean, D.M., 1966. Observations on halogens as bathing water disinfectants. Journal of Applied Bacteriology 29, 559–565.

Costerton, J.W., Stewart, P.S., Greenberg, E.P., 1999. Bacterial biofilms: a common cause of persistent infections. Science 284, 1318–1322.

Crandall, R.A., MacKenzie, C.J.G., 1984. Pathogenic hazards and public spa and hot tub facilities. Canadian Journal of Public Health 75, 223–226.

EPA, 1979. Efficacy Data Requirements: Swimming Pool Water Disinfectants. . April 23, DIS/TSS-12 www.epa.gov/oppad001/dis_tss_docs/dis-12.htm.

Fallon, R.J., Rowbotham, T.J., 1990. Microbiological investigations into an outbreak of Pontiac fever due to *Legionella micdadei* associated with use of a whirlpool. Journal of Clinical Pathology 43 (6), 479–483.

Favero, M.S., 1984. Whirlpool spa-associated infections: are we really in hot water? AJPH 74 (7), 653–655.

Fields, B.S., Haupt, T., Davis, J.P., 2001. Pontiac fever due to *Legionella micdadei* from a whirlpool spa: possible role of bacterial endotoxin. Journal of Infectious Diseases 184, 1289–1292.

Goeres, D.M., Palys, T., Sandel, B.B., Geiger, J., 2004. Evaluation of disinfectant efficacy against biofilm and suspended bacteria in a laboratory swimming pool model. Water Research 38, 3103–3109.

Goeres, D.G., Loetterle, L.R., Hamilton, M.A., Murga, R., Kirby, D.W., Donlan, R.M., 2005. Statistical assessment of a laboratory method for growing biofilms. Microbiology 151, 757–762.

Gustafson, T.L., Band, J.D., Hutcheson Jr., R.H., Schaffner, W., 1983. *Pseudomonas* folliculitis: an outbreak and review. Review of Infectious Diseases 5 (1), 1–8.

Hamilton, M.A., 2002. Testing antimicrobials against biofilm bacteria. Journal of AOAC International 85 (2), 479–485.

Havelaar, A.H., Bosman, M., Borst, J., 1983. Otitis externa by *Pseudomonas aeruginosa* associated with whirlpools. Journal of Hygiene Cambridge 90, 489–498.

Highsmith, A.K., Favero, M.S., 1985. Microbiologic aspects of public whirlpools. Clinical Microbiology Newsletter 7 (2), 9–11.

Highsmith, A.K., McNamara, M., 1988. Microbiology of recreational and therapeutic whirlpools. Toxicity Assessment 3, 599–611.

Holmes, S.E., Kinde, M.R., Person, J.L., Hennes, R.F., 1989. Gastroenteritis outbreak disease linked to swimming pool and spa use. Journal of Environmental Health 51 (5), 286–288.

Insler, M.S., Gore, H., 1986. *Pseudomonas* keratitis and folliculitis from whirlpool exposure. American Journal of Ophthalmology 101, 41–43.

IUPAC (International Union of Pure and Applied Chemistry), 1997. In: McNaught, A.D., Wilkinson, A. (Eds.), Compendium of Chemical Terminology, 2nd Edition. Blackwell Scientific Publications, Oxford. http://www.iupac.org/publications/compendium/index.html.

Judd, S.J., Black, S.H., 2000. Disinfection by-product formation in swimming pool waters: a simple mass balance. Water Research 34 (5), 1611–1619.

Kahana, L.M., Kay, J.M., Yakrun, M.A., Waserman, S., 1997. *Mycobacterium avium* complex infection in an immunocompetent young adult related to hot tub exposure. Chest 111 (1), 242–245.

Kim, H., Shim, J., Lee, S., 2002. Formation of disinfection by-products in chlorinated swimming pool water. Chemosphere 46, 123–130.

Koski, T.A., Stuart, L.S., Ortenzio, L.F., 1966. Comparison of chlorine, bromine, and iodine as disinfectants for swimming pool water. Applied Microbiology 14 (2), 276–279.

Kush, B.J., Hoadley, A.W., 1980. A preliminary survey of the association of *Pseudomonas aeruginosa* with commercial whirlpool bath waters. AJPH 70 (3), 279–281.

Lewis, K., 2001. Mini-review: riddle of biofilm resistance. Antimicrobial Agents and Chemotherapy 45 (4), 999–1007.

Mangione, E.J., Huitt, G., Lenaway, D., Beebe, J., Bailey, A., Figoski, M., Rau, M.P., Albrecht, K.D., Yakrus, M.A., 2001. Nontuberculous mycobacterial disease following hot tub exposure. Emerging Infectious Diseases 7 (6), 1039–1042.

McCausland, W.J., Cox, P.J., 1975. Pseudomonas infection traced to motel whirlpool. Journal of Environmental Health 37 (5), 455–459.

Murga, R., Forster, T.S., Brown, E., Pruckler, J.M., Fields, B.S., Donlan, R.M., 2001. Role of biofilms in the survival of *Legionella pneumophila* in a model potable-water system. Microbiology 147, 3121–3125.

Ortenzio, L.F., Clark, L.S., 1964. A standard test for efficacy of germicides and acceptability of residual disinfecting activity in swimming pool water. Journal of AOAC International 47 (3), 540–547.

Price, D., Ahearn, D.G., 1988. Incidence and persistence of *Pseudomonas aeruginosa* in whirlpools. Journal of Clinical Microbiology 26 (9), 1650–1654.

Primm, T.P., Lucero, C.A., Falkinham, J.O., 2004. Health impacts of environmental *Mycobacteria*. Clinical Microbiology Reviews 17 (1), 98–106.

Rickman, O.B., Ryu, J.H., Fidler, M.E., Kalra, S., 2002. Hypersensitivity pneumonitis associated with *Mycobacterium avium* complex and hot tub use. Mayo Clinic Proceedings 77, 1233–1237.

Salmen, P., Dwyer, D.M., Vorse, H., Kruse, W., 1983. Whirlpool-associated *Pseudomonas aeruginosa* urinary tract infections. JAMA 250 (15), 2025–2026.

Samples, J.R., Binder, P.S., Luibel, F.J., Font, R.L., 1984. *Acanthamoeba* keratitis possibly acquired from a hot tub. Archives of Ophthalmology 102, 707–710.

Shaw, J.W., 1984. A retrospective comparison of the effectiveness of bromination and chlorination in controlling *Pseudomonas aeruginosa* in spas (whirlpools) in Alberta. Canadian Journal of Public Health 75, 61–68.

Spitalny, K.C., Vogt, R.L., Orciari, L.A., Witherell, L.E., Etkind, P., Novick, L.F., 1984a. Pontiac fever associated with a whirlpool spa. American Journal of Epidemiology 120 (6), 809–817.

Spitalny, K.C., Vogt, R.L., Witherell, L.E., 1984b. National survey on outbreaks associated with whirlpool spas. AJPH 74 (7), 725–726.

Thomas, D.L, Mundy, L.M., Tucker, P.C., 1993. Hot tub legionellosis. Archives of Internal Medicine 153, 2597–2599.

Tilt, N., Hamilton, M.A., 1999. Repeatability and reproducibility of germicide tests: a literature review. Journal of AOAC International 82 (2), 384–389.

Wright, J.B., 2000. *Legionella* biofilms: their implications, study and control. In: Evans, L.V. (Ed.), Biofilms: Recent Advances in Their Study and Control. Harwood Academic Publishers, The Netherlands, pp. 291–310.

Zelver, N., Hamilton, M., Goeres, D., Heersink, J., 2001. Development of a standardized antibiofilm test. In: Doyle, R.J. (Ed.), Methods in Enzymology—Biofilms II, vol. 337. Academic Press, San Diego, CA, pp. 363–376.