MICROBIAL RISK ASSESSMENT OF GRAYWATER REUSE USING QUANTITATIVE MOLECULAR APPROACHES

By
Francis L. de los Reyes III*, H. Christopher Frey, Jory Whalen

Department of Civil, Construction and Environmental Engineering
NC State University
Raleigh, NC

*Lead Principal Investigator
The research on which this report is based was supported by funds provided by the Urban Water Consortium through the Water Resources Research Institute.

Contents of this publication do not necessarily reflect the views and policies of the UWC or of WRRI, nor does mention of trade names or commercial products constitute their endorsement by the WRRI or the State of North Carolina.

This report fulfills the requirements for a project completion report of the Water Resources Research Institute of The University of North Carolina. The authors are solely responsible for the content and completeness of the report.

WRRI Project No. 11-04-U
January 2014
ABSTRACT

Graywater is wastewater from bathing and washing facilities that does not contain concentrated human waste (toilet wastewater). About 50-75% of the total water used in homes results in graywater, and appropriate reuse of graywater can reduce the amount of treated water needed in the home. However, uncertainty regarding the risks associated with graywater reuse has prevented more substantial use of graywater. The objective of this study was to use molecular approaches to quantify pathogens present in graywater samples and to assess the risks associated with graywater reuse. Samples were collected from 30 households with different demographics. DNA was extracted from the collected graywater samples, and the presence and concentration of pathogens were determined using qPCR targeting specific pathogen DNA. A quantitative microbial risk assessment based on pathogen concentrations determined using molecular techniques was performed, and compared to risk assessment using fecal indicators. We also report the first next-generation sequencing of DNA from graywater. Finally, an experiment was conducted to determine if storage of untreated graywater affects the growth of bacteria.

Of the graywater samples collected from the different households, no samples tested positive for *C. jejuni*, one out of eight tested positive for *L. pneumophila*, and two out of eight tested positive for *A. hydrophila*. The results show that pathogen contamination occurs less frequently (i.e., in very few households) than estimated by models using fecal indicators to estimate pathogen concentration. However, when pathogens are detected, the risk of infection, as determined by a quantitative microbial risk assessment, exceeds the risk estimated by models that used fecal indicators. The metagenomic DNA sequencing results show that the microbial communities are significantly different for each source of graywater, indicating the human health risks associated with reuse could vary from source to source (i.e., shower/bath, laundry, bathroom sink). Graywater samples stored for different periods showed an initial increase in the concentration of bacteria followed by stabilization in bacterial concentration as organic matter is depleted. Taken together, the results have significant implications for future studies and policy. First, the use of fecal indicators for risk assessment of graywater reuse leads to highly variable and inaccurate results. Thus, future risk assessment should take into account concentrations of viable pathogens, and molecular methods that quantify viable pathogen levels should be pursued. In these studies, analyzing the pathogen levels from different sources may be needed, and will impact reuse options for different sources of graywater. Combining graywater from different households may lead to more frequent detection of pathogens, and individual graywater collection systems will result to lower risks for a greater number of households. Graywater treatment at the household level and lowering exposure during reuse are needed to lower the risks of infection due to pathogens. Future studies assessing treatment options are recommended. Finally, it is recommended that raw graywater water not be stored for more than 24 hours.
ACKNOWLEDGMENTS

We wish to thank all the participants of the survey, who allowed us to collect graywater samples from their homes. We also thank Dr. Jorge W. Santo Domingo, Dr. Hodon Ryu, and Dr. Vicente Gomez-Alvarez from the U.S. Environmental Protection Agency for their assistance with genomic sequencing. We also gratefully acknowledge the support and assistance of Jason So, Kara Kopf, and David Black of the NCSU Environmental Engineering Laboratory.
1. Introduction

Graywater is a valuable resource that can be utilized to help meet increasing water demands. Graywater is defined as wastewater from bathing and washing facilities that does not contain concentrated human waste (i.e., flush water from toilets) or food waste (i.e., kitchen sink, food waste grinders) (Asano et al., 2006). The absence of human and food waste implies that graywater is relatively free of pathogens and organic matter. The concentration of pathogens in graywater is typically much lower than in municipal wastewater. Fecal indicators such as total coliforms and fecal coliforms are generally lower in graywater than in wastewater, as shown in Table 1, suggesting that the pathogen load of graywater is less.

Table 1. Fecal indicators in graywater and wastewater

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Concentration in raw residential wastewater (cells per ml)</th>
<th>Concentration in raw residential graywater (cells per ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Coliform</td>
<td>$1.0 \times 10^8$ to $1.0 \times 10^{10}$</td>
<td>$5.0 \times 10^2$ to $2.4 \times 10^8$</td>
</tr>
<tr>
<td>Fecal Coliform</td>
<td>$1.0 \times 10^6$ to $1.0 \times 10^8$</td>
<td>$3.2 \times 10^4$ to $7.9 \times 10^6$</td>
</tr>
</tbody>
</table>

(Christova-Boal et al. 1996; Otis et al. 2002; Rose et al. 1991; Winward et al. 2008)

The average indoor water use for North America is 69 gallons per capita per day (Mayer and DeOreo 1999). About 50 to 75% of the indoor water used in the home contributes to graywater (Maimon et al. 2010). Based on these assumptions the average household produces between 34.5 and 51.8 gallons of graywater per capita per day.

Figure 1. Indoor water use in North America (Mayer and DeOreo 1999)
Due to the lower pathogens load and the abundance of graywater there is great potential for implementing reuse systems that separate graywater from other wastewater sources. Separated graywater could be reused for end-uses that do not require highly treated water. These non-potable end-uses include irrigation, toilet flushing, and vehicle and home washing, among others. Reusing graywater will reduce the amount of potable water that is needed in the home. Additionally, separating graywater from wastewater will reduce the amount of wastewater being discharged into the existing wastewater distribution system, which will also reduce the amount of wastewater that needs to be treated (Maimon et al. 2010).

Despite the potential benefits there remains hesitancy among municipalities and the general public to reuse graywater. Only 7% of households in the U.S. are reusing graywater (Roesner et al. 2006). This is likely due to unclear human health risks from exposure to untreated graywater. Currently there are no documented cases of public health impacts from exposure to untreated graywater. However, there is inadequate information regarding the human health risks from exposure to untreated graywater (Sheikh 2010).

Graywater may contain pathogenic organisms, which may lead to unacceptably high risks. There is very limited information regarding the microbial characteristics of graywater. The few reports that have sought to quantify pathogens in graywater have used traditional culture based methods and fecal indicators to estimate the concentration of the pathogens. The limitations of using these methods are that not all organisms are targeted and fecal indicators do not always correspond with the actual pathogens.

A few reports have sought to utilize fecal indicators and traditional culture methods to estimate the human health risks associated with graywater reuse. These reports have relied on indicators such as thermal tolerant coliforms, enterocci and Eschericia coli to quantify pathogens such as Salmonella, and rotaviruses (Godfrey et al. 2009; Maimon et al. 2010; Ottoson and Stenstrom 2003). Ottoson and Stenstrom (2003) measured common chemical biomarkers coprostanol and cholesterol in graywater samples to estimate the pathogen concentration of Campylobacter jejuni, Salmonella, Cryptosporidium parvum, Giardia lamblia and rotaviruses. These reports defined exposure scenarios such as accidental ingestion, indirect ingestion from hand-to-mouth contact from plants and lawns irrigated with graywater, ingestion of sprays from graywater irrigation systems and ingestion of soil contaminated with graywater. Using the pathogen concentrations and exposure scenarios, the number of pathogens, or pathogen dose that an individual ingests from reusing graywater was estimated. After the dose was determined, the risk of infection was calculated using exponential or beta-Poisson distributions. The reported risks were presented as risks of infection, and disability adjusted life years (DALY).

Ottoson and Stenstrom (2003) determined the greatest risk of reusing graywater was due to infection by rotavirus, which represented an unacceptably high risk of infection between 0.04 and 0.60 for the defined scenarios. Maimon et al. (2010) stated that according to the maximum tolerable risk for waterborne disease set by the World Health Organization (WHO, 2006), a safe pathogen dose would be $1.4 \times 10^{-4}$ rotaviruses/mL which would suggest a maximum tolerable concentration of $E. coli$ between $10^2$ to $10^3$ per 100 mL . Godfrey et al.
(2009) determined that based on the defined scenario, the risks of reusing graywater were between $0.2 \times 10^{-6}$ and $6.0 \times 10^{-6}$ DALY per year.

These studies report risk and provide a framework for an initial quantitative microbial risk assessment (QMRA) of graywater reuse. However, limitations to their estimations of the human health risks from exposure to untreated graywater remain. One of the major limitations of these estimates of risk is the uncertainty regarding the concentration of pathogens in graywater. In each of these models, indicators were used to predict the concentration of different waterborne pathogens, which is known to be an inaccurate method of estimating pathogen concentrations. The purpose of this study is to use molecular approaches to quantify key pathogens in graywater to analyze the human health risks associated with reusing untreated graywater.

2. Methods

The overall objective of the project is to quantify the pathogens in graywater and perform an initial quantitative microbial risk assessment of graywater (GW) reuse. The specific objectives are: (1) To quantify the levels of bacterial and protozoan pathogens in graywater using quantitative PCR; and (2) To perform an initial microbial risk assessment that includes: (i) quantifying the exposure to pathogens; and (ii) using the appropriate dose-response curves for each detected pathogen to estimate the risk of infection. A third objective was included in the original proposed work: to determine the fate of pathogens in soil irrigated with untreated graywater and in toilets flushed with gray water. The data from the pathogen quantification and risk assessment showed that treatment was needed even at collection, given the high risks associated with graywater reuse. Thus performing the experiments to meet the third objective was not warranted, since treating graywater is the recommendation. The additional time and resources to perform the experiments made performing these experiments not possible given the constraints of the project. An additional study on the effect of storage time on untreated graywater was deemed to be more relevant, and was performed.

The completed tasks include (1) optimized sampling and DNA extraction methods, (2) collection of samples from various households, (3) quantification of pathogens in graywater using Quantitative Polymerase Chain Reaction, (4) development of a preliminary risk assessment model to estimate the risks of reusing graywater, (5) comparison with an updated risk assessment model using the molecular analysis results, (6) next generation sequencing (pyrosequencing) to complete the microbial community assessment of graywater samples, and (7) monitoring of bacterial changes during storage of raw graywater. Next generation sequencing was performed in collaboration with the US EPA. We are continuing the study beyond the scope of the original project by performing additional Illumina sequencing with our collaborators from US EPA. This work was affected by the federal government shutdown, but is ongoing.
2.1. Sample Collection

Graywater samples were collected from 30 households in Wake County, North Carolina. When possible each household provided a sample from three sources: (1) Shower/Bath; (2) Bathroom Sink; and (3) Laundry. These samples were collected using 250mL and 500 mL sterilized bottles. From the 30 households 29 shower/bath samples were collected, 28 bathroom sink samples were collected and 23 laundry samples were collected. There was a great deal of variability with respect to the activities contributing to graywater samples. For example, some homes contributed bathroom sink samples that included teeth brushing and hand washing while others may have included shaving and washing faces. Several of these factors where tracked by having participants fill out a brief survey regarding the collected samples. This survey is shown on the following page.
## Participant Survey

### General Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes □ No □</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laundry Sample?</td>
<td></td>
</tr>
<tr>
<td>Bathroom Sink?</td>
<td></td>
</tr>
<tr>
<td>Bath/Shower?</td>
<td></td>
</tr>
<tr>
<td>Household Size:</td>
<td></td>
</tr>
<tr>
<td>Number of Adult Males:</td>
<td></td>
</tr>
<tr>
<td>Number of Adult Females:</td>
<td></td>
</tr>
<tr>
<td>Number of Children:</td>
<td></td>
</tr>
<tr>
<td>Number of Pets:</td>
<td></td>
</tr>
<tr>
<td>Type of Pets:</td>
<td></td>
</tr>
</tbody>
</table>

### Bathroom Sink Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes □ No □</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Individuals Contributing to Bathroom Sink Sample:</td>
<td></td>
</tr>
<tr>
<td>Did Children Contribute to Bathroom Sink Sample?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Teeth Brushing?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Hand Washing?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Face Washing?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Shaving?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Other Considerations Regarding Bathroom Sink Sample:</td>
<td></td>
</tr>
</tbody>
</table>

### Laundry Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes □ No □</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wash Cycle?</td>
<td></td>
</tr>
<tr>
<td>Rinse Cycle?</td>
<td></td>
</tr>
<tr>
<td>Detergent?</td>
<td></td>
</tr>
<tr>
<td>Bleach?</td>
<td></td>
</tr>
<tr>
<td>Did Children Contribute to Laundry:</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Number of Individuals Contributing to Laundry:</td>
<td></td>
</tr>
<tr>
<td>Other Considerations Regarding Laundry Sample:</td>
<td></td>
</tr>
</tbody>
</table>

### Bath/Shower Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes □ No □</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Individuals Contributing to Bath/Shower Sample:</td>
<td></td>
</tr>
<tr>
<td>Did Children Contribute to Bath/Shower Sample?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Bath?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Shower?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>Other Considerations Regarding Bath/Shower Sample:</td>
<td></td>
</tr>
</tbody>
</table>

---

8
2.2. Filtration and DNA Extraction Methods

Before samples could be analyzed, filtration and DNA extraction needed to be optimized. Initially, graywater samples were collected from one household for testing of processing and DNA extraction methods. Graywater samples were filtered so that the microorganisms could be concentrated and then the DNA extracted from the filter paper. Several filtration approaches were tested. These included using 0.2 \( \mu \text{m} \) filter paper and a vacuum pump (Figure 2a), using 0.4 \( \mu \text{m} \) filter paper and using inline cartridge filters and a peristaltic pump (Figure 2b).

![Figure 2. (a) Thermo Scientific Nalgene® Polysulfone Filter Holder; (b) Sterivex Inline Cartridge Filter](image)

Both 0.22\( \mu \text{m} \) and 0.45\( \mu \text{m} \) Sterivex filters were analyzed. To determine the feasibility of using Sterivex filters, filtering times and volumes were compared for each filtering method.

Samples were collected, transported on ice to the NCSU Environmental Engineering Laboratory, and filtered within 24 hours of collection. Based on the results of the initial study, subsequent samples were filtered using 0.45 \( \mu \text{m} \) Sterivex filters and a peristaltic pump.

Filter paper was then removed from Sterivex filters and DNA was extracted from filter paper using the Aluminum Sulfate method developed in the NCSU Environmental Engineering Laboratory.

2.3. Quantitative Polymerase Chain Reaction (qPCR)

After the DNA was extracted from each of the graywater samples, qPCR was performed to detect and quantify targeted pathogens in graywater samples. Quantitative PCR uses specific primers and cycling parameters to amplify targeted DNA through a repeated process of denaturation, annealing and elongation. In qPCR, a fluorescent dye attaches to the DNA and the intensity of the fluorescent dye increases as the DNA concentrations increase during amplification. The increasing fluorescence intensity is tracked and reported using a \( C_t \) value which indicates the number of PCR cycles required to amplify the DNA to a designated
threshold. Standard curves are created plotting $C_t$ values versus known concentrations of DNA.

Three waterborne pathogens were targeted: Legionella pneumophila, Campylobacter jejuni, and Aeromonas hydrophila. In addition, Bacteroides was targeted as a fecal indicator. For the samples that tested positive for any of the four targeted organisms, a general bacterial primer was also used to quantify the bacteria in the samples.

DNA from *A. hydrophila* ATCC 7966D-5, *C. jejuni* ATCC 700819D-5, *L. pneumophila* ATCC 33152D-5, and *Bacteroides vulgatus* ATCC 8482D-5 were purchased from the American Type Culture Collection (ATCC). Additionally a standard curve was developed for the general bacterial primers using *Bacteroides vulgatus* ATCC 8482D-5.

Tenfold dilution series ranging from 1 ng/μL to $1 \times 10^{-9}$ ng/μL for *A. hydrophila*, 5 ng/μL to $5 \times 10^{-9}$ ng/μL for *C. jejuni*, 50 ng/μL to $5 \times 10^{-6}$ ng/μL, and 50 ng/μL to $5 \times 10^{-4}$ for *Bacteroides* were prepared. Quantitative PCR was then performed on triplicates of the dilution series using the Bio-Rad iQ5 iCycler Multicolor Real-Time PCR Detection System.

The total volume of the each qPCR reaction was 25 μL. This includes 7 μL sterile water, 0.25 μL of both the forward and reverse primers (25 μM), 12.5 μL SYBR Green mix (Bio-Rad iQ SYBR Green Supermix), and 5 μL of the DNA template. Primers and cycling parameters for each of the targeted organisms are shown in Table 2.

Table 2. Primers and cycling parameters for targeted organisms. (adapted from Ahmed et al., 2008 and Harms et al., 2003)

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer Sequence (5'-3')</th>
<th>PCR Cycling Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. hydrophila</em> lip gene</td>
<td>AAC CTG GTT CCG CTC AAG CCG TTG (F), TTG CTC GCC TCG GCC CAG CAG CT (R)</td>
<td>15 min at 95°C, 35 cycles of 60 s at 94°C, 60 s at 62°C, and 90 s at 72°C</td>
</tr>
<tr>
<td><em>C. jejuni</em> mapA gene</td>
<td>GCT AGA GGA ATA GTT GTG CTT GAC AA (F), TTA CTC ACA TAA GGT GAA TTT TGA TCG (R)</td>
<td>10 min at 95°C, 50 cycles of 15 s at 95°C and 30 s at 59°C</td>
</tr>
<tr>
<td><em>L. pneumophila</em> mip gene</td>
<td>GCA ATG TCA ACA GCAA (F), CAT AGC GTC TTG CATG (R)</td>
<td>15 min at 95°C, 35 cycles of 30 s at 94°C, 60 s at 54°C, and 60 s at 72°C</td>
</tr>
<tr>
<td><em>Bacteroides</em> 16S rRNA gene</td>
<td>AAC GCT AGC TAC AGG CTT (F), CAA TCG GAG TTC TTT GTG (R)</td>
<td>15 min at 95°C, 35 cycles of 30 s at 94°C, 60 s at 59°C, and 60 s at 72°C</td>
</tr>
</tbody>
</table>
Initially a single qPCR reaction was performed for each sample to determine the presence or absence of the targeted organism. qPCR analysis was then repeated to quantify pathogens in the samples that tested positive for the targeted pathogen. qPCR was performed in triplicate for each positive sample.

$C_t$ values for the serial dilutions were determined and a standard curve plotting the $C_t$ value versus the log concentration (ng/µL) of the template DNA was created for each pair of primers. A linear trendline was fit to the data to represent the relationship between the DNA concentration and the $C_t$ values. The $R^2$ and the PCR efficiencies were measured for each standard curve. The PCR efficiency represents how efficiently the DNA replicates. Ideally each cycle of qPCR will result in the DNA doubling. As a result, a 10-fold increase should occur after 3.32 cycles. Thus for a standard curve featuring 10-fold dilutions, the $C_t$ value should increase by 3.32 for each dilution, resulting in a slope of -3.32. The PCR efficiency can be calculated using Equation 1.

**Equation 1.** \[ \text{PCR Efficiency} = 10^{1/\text{slope} - 1} \]

Thus for perfect doubling the PCR efficiency is 100%. The $C_t$ values were then converted to a concentration (ng/µL), which was subsequently converted to cells/mL. Equation 2 was used to determine the mass of the DNA of each organism.

**Equation 2.**

\[
\text{DNA (fg)} = \left( \text{genome size (bp)} \right) \times \left( \frac{\text{MW (Da)}}{\text{bp}} \right) \times \left( \frac{1.66 \times 10^{-27} \text{ kg}}{\text{Da}} \right) \times \left( \frac{10^{18} \text{ fg}}{\text{kg}} \right)
\]

$\text{MW}$ represents the molecular weight in Daltons. The relative weight of a base pair is 615.3830 for AT pairs and 616.3711 for GC pairs. The average molecular weight can be calculated using Equation 3.

**Equation 3**

\[
\text{MW (Da/bp)} = (\text{GC content} \%) \times 616.3711 + (\text{AT content} \%) \left( \frac{\text{MW (Da)}}{\text{bp}} \right) \times 615.3830
\]

The genome size, GC content, and the AT content are shown in Table 3.
Table 3. Genome characteristics of targeted organisms

<table>
<thead>
<tr>
<th>Organism</th>
<th>Genome Size (bp)</th>
<th>GC Content</th>
<th>AT Content</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. hydrophila</em> ATCC 7966</td>
<td>4,744,448</td>
<td>61.5%</td>
<td>38.5%</td>
<td>Seshadri et al. (2006)</td>
</tr>
<tr>
<td><em>C. jejuni</em> ATCC 700819</td>
<td>1,641,481</td>
<td>30.6%</td>
<td>69.4%</td>
<td>Parkhill et al. (2000)</td>
</tr>
<tr>
<td><em>L. pneumophila</em> ATCC 33152</td>
<td>3,353,636</td>
<td>38.4%</td>
<td>61.6%</td>
<td>NCBI (2013)</td>
</tr>
<tr>
<td><em>Bacteroides vulgatus</em> ATCC 8482</td>
<td>4,781,702</td>
<td>42.0%</td>
<td>58.0%</td>
<td>Cuiv et al. (2011)</td>
</tr>
</tbody>
</table>

Note: Values shown for *L. pneumophila* represent the average value from several studies that have analyzed the genome sequences of various strands of *L. pneumophila*.

Using Equations 2 and 3 the relative mass of DNA of each organism was found (Table 4).

Table 4. Mass of targeted organisms

<table>
<thead>
<tr>
<th>Organism</th>
<th>Mass of DNA per Cell (fg/cell)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. hydrophila</em> ATCC 7966</td>
<td>4.85</td>
</tr>
<tr>
<td><em>C. jejuni</em> ATCC 700819</td>
<td>1.68</td>
</tr>
<tr>
<td><em>L. pneumophila</em> ATCC 33152</td>
<td>3.43</td>
</tr>
<tr>
<td><em>Bacteroides vulgatus</em> ATCC 8482</td>
<td>4.89</td>
</tr>
</tbody>
</table>

The number of cells per mL in each reaction was then calculated by taking the concentration (ng/μL) as determined by qPCR and dividing that value by the mass per cell as shown in Table 4. The number of cells per mL in the collected graywater samples was then determined using Equation 4.

**Equation 4**

\[
\frac{\text{cells}}{\text{mL}} = \left( \frac{\text{DNA concentration (ng/μL)}}{\text{Extract}} \right) \times \left( \frac{100 \, \text{μL}}{\text{Extract}} \right) \times \left( \frac{\text{Volume Filtered (mL)}}{\text{ng}} \right) \times \left( \frac{10^6 \, \text{fg}}{\text{Cell}} \right) \times \left( \frac{\text{Mass of Cell (fg)}}{\text{Cell}} \right)
\]
2.4. Preliminary Risk Assessment Model

A preliminary risk assessment model was used to evaluate the effect of using different fecal indicators to estimate pathogen concentrations and the risk of infection. The initial risk assessment model evaluated the risk of infection from exposure to a single pathogen (*Campylobacter jejuni*) in untreated graywater.

The risk assessment was based on the model used by Ottoson and Stenstrom (2003). The model consisted of three distinct parts: 1) the pathogen concentration model as shown in Figure 3; 2) the pathogen dose model as shown in Figure 4; and 3) the dose response models.

\[
FL = \frac{I}{ED}
\]

**Figure 3. Pathogen Concentration Model as explained by Ottoson and Stenstrom (2003)**

The fecal indicators used in the model by Ottoson and Stenstrom (2003) were in units of cfu (coliform forming units) per person per day. However, the values for the fecal indicators used in this study were obtained from other published reports. These studies quantified fecal indicators in units of cfu per mL of graywater. This resulted in a slight modification of the pathogen concentration model (Figure 4).
Figure 4. Modified Pathogen Concentration Model

The parameters shown in black represent the characteristics of the fecal indicators. The parameters shown in gray represent the characteristics of the pathogens. As shown in the model, fecal indicators represent the concentration of the indicators of fecal contamination in graywater. The excretion density represents the estimated concentration of each fecal indicator in human feces. Yearly incidence is the percentage of individuals who become infected each year. Excretion time is the duration that an infected individual will excrete the specified pathogen. Pathogen density is the estimated number of pathogens excreted.

Four different fecal indicators, *E. coli*, fecal coliforms, total coliforms, and *Enterococci*, were used to estimate the concentration of the pathogen *C. jejuni* in untreated graywater. A wide range of values for the concentration of fecal indicator concentration has been reported. However, there is no obvious distribution for the concentrations of the four fecal indicators that were used in the model. Therefore the concentrations of the four fecal indicators were represented by uniform distributions based on maximum and minimum values from several previously published studies as shown in Table 5. Data for the excretion density was obtained from previously published studies and is also shown in Table 5. When a range of values was reported for the excretion density, a uniform distribution was used because there was no obvious distribution.
Table 5. Fecal Indicator Characteristics

<table>
<thead>
<tr>
<th>Fecal Indicator</th>
<th>Fecal Indicators in Graywater (cfu/ml)(^a)^(^b)</th>
<th>Excretion Density (cfu/g feces)(^c)^(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>0 - 2.4 x 10(^4)</td>
<td>1.0 x 10(^7)</td>
</tr>
<tr>
<td>Total Coliforms</td>
<td>2.4 x 10(^2) - 2.4 x 10(^5)</td>
<td>1.0 x 10(^7) - 1.0 x 10(^9)</td>
</tr>
<tr>
<td>Fecal Coliforms</td>
<td>1.7 - 3.3 x 10(^3)</td>
<td>1.0 x 10(^7) - 1.0 x 10(^6)</td>
</tr>
<tr>
<td>Enterococci</td>
<td>0.0 - 2.0 x 10(^2)</td>
<td>3.2 x 10(^9)</td>
</tr>
</tbody>
</table>

\(^a\) (Birks and Hills, 2007)  
\(^b\) (Christova-Boal et al., 1996)  
\(^c\) (Geldreich, 1978)  
\(^d\) (Ottoson and Stenstrom, 2003)

Additionally two chemical biomarkers, coprostanol and cholesterol, were used as indicators to estimate the concentration of the pathogen *C. jejuni* in untreated graywater using the model shown in Figure 5.

\[
FL = \frac{CB}{ED}
\]

**Figure 5. Modified Pathogen Concentration Model using Chemical Biomarkers**

Figure 5 represents the same model as shown in Figure 4 except that chemical biomarkers were used to estimate the fecal load instead of fecal indicators.

The concentrations of the two biomarkers were represented by lognormal distributions as determined by Ottoson and Stenstrom (2003) (Table 6).
Table 6. Chemical Biomarker Characteristics (Ottoson and Stenstrom, 2003)

<table>
<thead>
<tr>
<th>Chemical Biomarker</th>
<th>Mean Biomarker Concentration (μg/L)</th>
<th>Standard Deviation (μg/L)</th>
<th>Excretion Density (mg/g feces)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coprostanol</td>
<td>8.6</td>
<td>4.4</td>
<td>12.74</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>17.3</td>
<td>8.4</td>
<td>5.08</td>
</tr>
</tbody>
</table>

Table 7 shows the other parameters that were used to estimate the concentration of *C. jejuni* in graywater.

Table 7. *C. jejuni* Characteristics (adapted from Ottoson and Stenstrom, 2003)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yearly Incidence (YI) (%)</td>
<td>15.6%</td>
</tr>
<tr>
<td>Excretion Time (ET) (Days)</td>
<td>μ = 16.14, σ = 2.11</td>
</tr>
<tr>
<td>Pathogen Density (PD) (#/g feces)</td>
<td>μ = 10^8, σ = 10</td>
</tr>
</tbody>
</table>

A summary of the values that were input into the modified pathogen concentration model is shown in Table 8.
Table 8. Summary of inputs for modified pathogen concentration model

<table>
<thead>
<tr>
<th>Input</th>
<th>Distribution</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fecal Indicators</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em> (CFU/mL)</td>
<td>Uniform (0, 2.4 x 10^4)</td>
<td>Birks and Hills (2007); Christova-Boal et al (1996)</td>
</tr>
<tr>
<td>Total Coliforms (CFU/mL)</td>
<td>Uniform (2.4 x 10^2, 2.4 x 10^5)</td>
<td>Birks and Hills (2007); Christova-Boal et al (1996)</td>
</tr>
<tr>
<td>Fecal Coliforms (CFU/mL)</td>
<td>Uniform (1.7, 3.3 x 10^3)</td>
<td>Birks and Hills (2007); Christova-Boal et al (1996)</td>
</tr>
<tr>
<td>Enterococci (CFU/mL)</td>
<td>Uniform (0, 2.2 x 10^5)</td>
<td>Birks and Hills (2007); Christova-Boal et al (1996)</td>
</tr>
<tr>
<td>Coprostanol (μg/L)</td>
<td>Lognormal(8.6, 4.4)</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
<tr>
<td>Cholesterol (μg/L)</td>
<td>Lognormal(17.3, 8.4)</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
<tr>
<td><strong>Excretion Density</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em> (CFU/g feces)</td>
<td>1.0 x 10^7</td>
<td>Geldreich (1978)</td>
</tr>
<tr>
<td>Total Coliforms (CFU/g feces)</td>
<td>Uniform (1.0 x 10^7, 1.0 x 10^9)</td>
<td>Geldreich (1978)</td>
</tr>
<tr>
<td>Fecal Coliforms  (CFU/g feces)</td>
<td>Uniform (1.0 x 10^5, 1.0 x 10^6)</td>
<td>Geldreich (1978)</td>
</tr>
<tr>
<td>Enterococci  (CFU/g feces)</td>
<td>Uniform (1.7, 3.3 x 10^5)</td>
<td>Geldreich (1978)</td>
</tr>
<tr>
<td>Coprostanol (mg/g feces)</td>
<td>12.74</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
<tr>
<td>Cholesterol (mg/g feces)</td>
<td>5.08</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
<tr>
<td><strong>Yearly Incidence (%)</strong></td>
<td>15.6%</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
<tr>
<td><strong>Pathogen Density (pathogens/g feces)</strong></td>
<td>Lognormal(10^8, 10)</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
</tbody>
</table>
Once the pathogen concentration was determined, an exposure assessment was performed to estimate the dose, or the number of pathogens an individual would consume. This exposure assessment is shown in Figure 6.

![Exposure Assessment Model](image)

**Figure 6. Exposure Assessment Model**

The ingested volume indicates the volume that a person would ingest during a given exposure scenario and the exposure frequency represents the number of exposures per year. Both the ingested volume and the exposure frequency are the result of four different exposure scenarios that were evaluated for this report. These exposure scenarios include accidental ingestion of graywater, hand-to-mouth contact from plants and lawns irrigated by graywater, ingestion of sprays used for irrigation and washing, and the consumption of crops irrigated with graywater. The ingested volumes and exposure frequencies from previous reports (Maimon et al., 2010, Tanaka et al., 1998) are shown in Table 9 for each exposure scenario.
Table 9. Exposure Scenarios

<table>
<thead>
<tr>
<th>Exposure Scenario</th>
<th>Description a,b</th>
<th>Ingested Volume (mL) a,b</th>
<th>Ingested Volume Range</th>
<th>Exposure Frequency a,b</th>
<th>Ingestion Frequency Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Accidental Ingestion of Graywater</td>
<td>100</td>
<td>50 - 200</td>
<td>1 per year</td>
<td>0.5 to 2 per year</td>
</tr>
<tr>
<td>2</td>
<td>Indirect Ingestion From Touching Plants and Lawns</td>
<td>1</td>
<td>0.5 – 2.0</td>
<td>90 per year</td>
<td>45 to 180 per year</td>
</tr>
<tr>
<td>3</td>
<td>Ingestion of Graywater From Sprays used for irrigation and washing</td>
<td>0.1</td>
<td>0.05 – 0.2</td>
<td>90 per year</td>
<td>45 to 180 per year</td>
</tr>
<tr>
<td>4</td>
<td>Ingestion of Crops Irrigated by Graywater</td>
<td>10</td>
<td>5 - 20</td>
<td>365 per year</td>
<td>183 to 730 per year</td>
</tr>
</tbody>
</table>

a (Maimon et al., 2010)  
b (Tanaka et al., 1998)

To account for variability due to the ingested volume and the exposure frequency, an estimate of the range of each parameter was made. This information was used in the model as a triangular distribution making the assumption that the average also represented the most likely value.

After the pathogen dose was estimated, the risk of infection was estimated using a dose response model. A Beta-Poisson dose response model (Equation 5) was used to estimate the risk of infection from exposure to *C. jejuni* (N_{50} = 896, \( \alpha = 0.145 \)) (Ottoson and Stenstrom, 2003).

**Equation 5**

\[
P(\text{response}) = 1 - \left[ 1 + \text{dose} \left( \frac{1}{N_{50}} - 1 \right) \right]^{-\alpha}
\]

The dose response model estimates the probability that an individual will be infected for a given dose. All quantitative evaluations were modeled using probabilistic modeling software, Analytica of Lumina Decision Systems, Inc. Analytica utilized the Monte Carlo method with 1000 iterations.

The inputs for the pathogen dose model and the dose response model are shown in Table 10.
Table 10. Summary of inputs for modified pathogen dose model and dose response model

<table>
<thead>
<tr>
<th>Input</th>
<th>Distribution</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ingested Volume (mL)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 1 (Accidental Ingestion)</td>
<td>Triangular (50, 100, 200)</td>
<td>Maimon (2010)</td>
</tr>
<tr>
<td>Scenario 2 (Indirect Ingestion)</td>
<td>Triangular (0.5, 1, 2)</td>
<td>Maimon (2010)</td>
</tr>
<tr>
<td>Scenario 3 (Ingestion of Sprays)</td>
<td>Triangular (0.05, 0.1, 0.2)</td>
<td>Maimon (2010)</td>
</tr>
<tr>
<td><strong>Exposure Frequency (Year⁻¹)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 1 (Accidental Ingestion)</td>
<td>Triangular (50, 100, 200)</td>
<td>Maimon (2010)</td>
</tr>
<tr>
<td>Scenario 2 (Indirect Ingestion)</td>
<td>Triangular (0.5, 1, 2)</td>
<td>Maimon (2010)</td>
</tr>
<tr>
<td>Scenario 3 (Ingestion of Sprays)</td>
<td>Triangular (0.05, 0.1, 0.2)</td>
<td>Maimon (2010)</td>
</tr>
<tr>
<td><strong>Dose Response C. jejuni</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N₅₀</td>
<td>896</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
<tr>
<td>α</td>
<td>0.145</td>
<td>Ottoson and Stenstrom (2003)</td>
</tr>
</tbody>
</table>

The preliminary risk assessment provided a rough estimate of the risks of reusing graywater. One major source of uncertainty is the model that used fecal indicators to estimate the concentration of *C. jejuni* in untreated graywater. Uncertainty arises because many pathogens may not correlate with indicator organisms (Harwood et al., 2005).

To evaluate which parameters contributed most to the uncertainty in the estimates of risk, an importance analysis was performed. Importance analysis is an analytical tool that estimates the most important inputs of a given model. This method estimates how much each uncertain input contributes to the uncertainty in the output. From the four exposure scenarios and the six indicators that were evaluated, there were 24 distributions that were generated. Four of these distributions were evaluated for importance. The four distributions that were evaluated represented each of the exposure scenarios and four of the six fecal indicators.

### 2.5. Updated Risk Assessment

Using the pathogen concentrations from the qPCR results, the human health risks were evaluated again. Quantifying pathogens in graywater using qPCR reduces the uncertainty in the risk assessment model because it eliminates the need to use fecal indicators to estimate the concentration of pathogens in graywater.

The risk assessment model used in the preliminary risk analysis was used to estimate the risk of infection from exposure to the targeted pathogens in graywater. However, the model could be simplified because the pathogen concentrations were determined using qPCR analysis, allowing
for the pathogen concentration model shown in Figure 5 to be eliminated from the risk assessment model. The same exposure scenarios that were used in the preliminary risk assessment were used to estimate the pathogen dose (Figure 6).

The risks of infection from exposure to *C. jejuni* and *L. pneumophila* in graywater were both estimated. The risk of infection from exposure to *A. hydrophila* was not calculated because there is a lack of correlation between *A. hydrophila* and the pathogenicity for humans (Morgan et al., 1985). A Beta-Poisson dose response model (Equation 6) was used to estimate the risk of infection from exposure to *C. jejuni* ($N_{50} = 896$, $\alpha = 0.145$) (Ottoson and Stenstrom, 2003).

**Equation 6**

$$P(\text{response}) = 1 - \left[ 1 + \frac{\frac{1}{\alpha} - 1}{N_{50}} \right]^{-\alpha}$$

An exponential dose response curve (Equation 7) was used to estimate the risk of infection from *L. pneumophila* ($k = 5.99 \times 10^{-2}$) (Muller et al., 1983).

**Equation 7**

$$P(\text{response}) = 1 - e^{-k \times \text{dose}}$$

The dose response models estimate the probability that an individual will be infected for a given dose. A $P(\text{response})$ equal to one suggests that an individual will be infected and a $P(\text{response})$ equal to zero indicates that there is no risk of infection.

### 2.6. Microbial Community Analysis and Effect of Storage

#### 2.6.1. Description of Graywater Collected for Storage Experiment

Graywater samples were collected from five households near Raleigh, North Carolina. Three of the households that provided graywater samples had children and one household had a dog. Each household provided between 1 and 1.5 liters of graywater from the bathroom sink, laundry and shower or bath.

The total amount of graywater included approximately 5.75 L from the bathroom sink, 7.5 L from the laundry, and 6.25 L from the shower or bath for a total of approximately 19.5 L of graywater. Samples were combined to fill four plastic buckets. Each bucket contained just under 5 L of graywater. The first bucket contained only graywater from bathroom sinks, the second bucket contained graywater only from the laundry, the third bucket contained only graywater from the showers or baths, and the fourth bucket contained combined graywater from each of the previous sources mentioned. The stored graywater was covered but not sealed using a plastic lid. Samples were stored for a period of 23 days.
2.6.2. Sampling Procedures and Coliform Assays

Samples were collected from each bucket on days 1, 2, 4, 9, 16, and 23. Day 1 represents the day when samples were collected and placed in buckets for storage. Stored graywater was mixed gently prior to sampling and approximately 750 mL of graywater was collected at each sampling period.

Collected samples were analyzed using both traditional culture based methods and molecular techniques. Traditional culture based methods included total and fecal coliform assays. Molecular analysis included qPCR targeting several waterborne pathogens. A total coliform assay was performed using the membrane filtration method as described in Standard Methods (Eaton and Franson, 2005). A fecal coliform assay was prepared for each sample that was collected. The fecal coliform test was performed using the membrane filtration method as described in Standard Methods (Eaton and Franson, 2005).

2.6.3. Pyrosequencing

Extracted DNA samples were sent to the US EPA National Water Quality Laboratory in Cincinnati, Ohio for pyrosequencing. Metagenome libraries were generated with the 454 Life Sciences GS-FLX Titanium platform (Gomez-Alvarez et al., 2012). Pyrosequencing reads were used for metagenomic analyses to determine the distribution of members of the Bacteria domain present in the graywater samples.

Non-metric multidimensional scaling (nMDS) and cluster analysis were performed by EPA to identify the relationships between the different community structures in the different graywater samples. This comparison was assessed using a Bray-Curtis similarity coefficient calculated from the relative distributions of the operational taxonomic units (OTU). Additionally, an analysis of similarity (ANOSIM) based on the Bray Curtis similarity matrix was conducted to determine the similarities between the different sources of graywater.
3. Results and Discussion

3.1. Sampling of Households

An attempt was made to represent a variety of households with respect to size, number of children and number of pets. Figures 7 and 8 show the diversity of the selected households with respect to pets and children. The average size of the households selected was 2.8 individuals per household.

![Figure 7. Number of Children per Household](image1)

![Figure 8. Number of Pets per Household](image2)

3.2 Optimizing Filtration and DNA Extraction

Several filtration methods were tested. The criteria for choosing the filtration method included: short filtration times, adequate DNA yield, and suitability for downstream PCR. The filtration methods tried included using 0.2 µm filter paper and a vacuum pump, using 0.4 µm filter paper, and using 0.22 µm and 0.45 µm inline cartridge filters and a peristaltic pump.

Using the 0.4 µm filter paper initially provided optimistic results as the first 250 mL were filtered within 3 minutes, however the pores soon became clogged and the next 150 mL took over 1 hour to filter. This indicated that to filter a volume of 500 mL, 0.4 µm filters and 0.2 µm filters would result in similar filtering times.

To determine the feasibility of using Sterivex filters, filtering times and volumes were compared for each filtering method. The filtering times were defined as the time taken until the filter clogged and filtration rate decreased to near zero.
The filtering times are much lower for the Sterivex filters, although the filtered volumes were similar. It was expected that as more samples are filtered, the range of the volumes filtered by each of these methods will be similar.

Due to shorter filtration times, the Sterivex filters were further analyzed to determine the DNA yield, quality, and success of PCR amplification. This analysis was based on the comparison of the same sample filtered with a 0.22 μm Sterivex filter and a 0.45 μm Sterivex filter as shown in Table 11.
Table 11. Comparison of 0.22 μm and 0.45 μm Sterivex Filters

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>0.22 μm Sterivex Filter</th>
<th>0.45 μm Sterivex Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Type</td>
<td>SH-4</td>
<td>SH-6</td>
</tr>
<tr>
<td>Volume (mL)</td>
<td>600</td>
<td>200</td>
</tr>
<tr>
<td>DNA Concentration (ng/μL)</td>
<td>11.59</td>
<td>5.59</td>
</tr>
</tbody>
</table>

Interestingly, the 0.22 μm Sterivex filtered a larger volume and yielded a higher concentration of DNA compared to the 0.45 μm filter. This was surprising but may be related to the higher initial flow rates in the larger filter, resulting in faster clogging. PCR was then performed on the extracted DNA using universal bacterial primers 8f and 1492r to determine the presence of bacteria in the filtered samples (Figure 11).

![Figure 11. PCR Results using a universal bacterial primer](image)

(A) Positive Control (5ng); (B) Positive Control (10ng); (C) SH-4 Shower Water 0.22μm filter (5ng); (D) SH-4 Shower Water 0.22μm filter (10ng); (E) SH-6 Shower Water 0.45μm filter (5ng); (F) SH-6 Shower Water 0.45μm filter (10ng); (G) Negative Control

All of the filtered samples showed a positive result as indicated in Figure 11. Considering all the factors of filtration time, DNA yield, and PCR amplification, it was decided that the 0.45 μm Sterivex filters should be used for filtering collected samples. While the 0.22 μm filters were
able to filter larger volumes and yield higher DNA concentrations, the 0.45 μm Sterivex filter resulted in a more distinct PCR product, as shown in Figure 11.

Samples from 30 households were collected and filtered. During the filtration process the filter pores would often clog, resulting in variability in the volume of sample that was filtered. A summary of the filtered volumes is shown in Figure 12.

![Figure 12. Filtered Volumes for Collected Samples](image)

The boxplots shown in Figure 12 show the median, upper quartile and lower quartile of the filtered volumes. The mean volume filtered is shown by the white diamond. The bars represent the 5th and 95th percentiles of the filtered volumes.

The mean filtered volumes were 197 mL, 209 mL, and 262 mL for samples collected from the bathroom sink, laundry, and shower/bath, respectively. The median volumes that were filtered were 175 mL, 175 mL, and 225 mL for samples collected from the bathroom sink, laundry, and shower/bath, respectively.
The boxplots shown in Figure 13 illustrate the median, upper quartile and lower quartile of the extracted DNA. The mean volume filtered is shown by the white diamond. The error bars represent the 5\textsuperscript{th} and 95\textsuperscript{th} percentiles of the extracted DNA.

The mean DNA concentrations extracted were 26.42 ng/μL, 28.95 ng/μL, and 17.75 ng/μL for samples collected from the bathroom sink, laundry, and shower/bath, respectively. The median DNA concentrations extracted were 17.41 ng/μL, 10.25 ng/μL, and 7.58 ng/μL for samples collected from the bathroom sink, laundry, and shower/bath, respectively.

3.3. Quantitative PCR

To quantify the selected pathogens, standard curves were developed (Figures 14 to 18). The standard curved show $R^2$ values of 0.99 or higher.
Figure 14. Standard Curve for *Aeromonas hydrophila* (PCR efficiency=0.87%)

\[ y = -3.6828x + 14.324 \]
\[ R^2 = 0.9937 \]

Figure 15. Standard Curve for *Campylobacter jejuni* (PCR efficiency=96%)

\[ y = -3.4156x + 15.297 \]
\[ R^2 = 0.9925 \]
Figure 16. Standard Curve for *Legionella pneumophila* (PCR efficiency=91%)

\[ y = -3.5474x + 14.841 \]
\[ R^2 = 0.9933 \]

Figure 17. Standard Curve for *Bacteroides* (PCR efficiency=80%)

\[ y = -3.9133x + 13.993 \]
\[ R^2 = 0.9975 \]
Figure 18. Standard Curve for Bacterial Primers using *Bacteroides* (PCR efficiency=92%)

Very few samples tested positive for the targeted organisms. A summary of the samples that tested positive for each of the targeted organisms is shown in Figure 19.

Five out of eighty samples tested positive for *Bacteroides*, two out of eighty samples tested positive for *Aeromonas hydrophila*, and one out of eighty samples tested positive for *Legionella pneumophila*. No samples tested positive for *Campylobacter jejuni*. Each of the two samples that tested positive for *Aeromonas hydrophila*, and the one sample that tested positive for
**Legionella pneumophila** also tested positive for *Bacteroides*. One household tested positive for *Bacteroides* in both their bathroom sink sample and their shower sample. Of the households that tested positive for *Bacteroides*, two households had children and two did not have any children. Each of the bathroom sinks that tested positive for *Bacteroides* included graywater that was resulted at least partially from the brushing of teeth.

Figure 20 shows a summary of the concentration of pathogens in the samples that tested positive for the targeted pathogens.

![Figure 20](image.png)

**Figure 7 Mean Pathogen Concentrations** Error Bars represent maximum and minimum concentrations, n = 3.

The average concentrations range from 29.3 to $1.63 \times 10^4$ cells per mL for *Bacteroides*, 8.7 to 24.3 cells per mL for *A. hydrophila* and 33.3 cells per mL for *L. pneumophila*. Concentrations are based on the presence of targeted DNA and do not account for the viability of the organisms. This suggests that these concentrations are conservative estimates of the actual viable organisms present in the graywater samples.

The houses that tested positive for the targeted organisms include households 2, 7 and 9. Household 2 included two adults (1 male, 1 female) and no children or pets. Only the bathroom sink graywater from household 2 tested positive for *A. hydrophila*, *L. pneumophila* and *Bacteroides*. The collected bathroom sink graywater from household 2 included graywater generated from hand washing, face washing, and brushing teeth.
There were five people in household 7, two adults and three children. Household 7 also had both a dog and a cat present in the home. Two graywater samples were collected from both the bathroom sink graywater and the shower/bath. One of the samples collected from the bathroom sink graywater tested positive for *A. hydrophila*. Both the bathroom sink samples and one of the shower/bath samples tested positive for *Bacteroides*. The collected bathroom sink graywater was generated from shaving and tooth brushing. The collected shower/bath graywater was generated from baths and showers of all five members of the household.

Household 9 included one adult male and no children or pets. Only the bathroom sink graywater from household 9 tested positive for *Bacteroides*. The collected bathroom sink graywater from household 9 was generated from hand washing, face washing, and brushing teeth.

The concentration of bacteria present in each of the samples is shown in Figure 21.

![Figure 21. Concentration of General Bacteria](image)

Note: (BS) indicates the bathroom sink samples and (SH) indicates the shower/bath samples.
Two (2) bathroom sink samples and two (2) shower/bath samples were collected from household 7.
Error bars represent the range.

The mean concentration of bacteria in the samples that tested positive for the other organisms ranged from $1.41 \times 10^5$ to $7.97 \times 10^6$ cells per mL. This indicates that the organisms targeted by qPCR represent only a small portion of the bacteria present in graywater samples. qPCR was not performed for Household 7 (BS-A) because there was no DNA remaining after previous qPCR assays were performed.
3.4. Preliminary Risk Assessment Model

The pathogen concentration model resulted in six distributions representing the concentration of *C. jejuni* in untreated graywater as estimated using the various indicators. The results of the pathogen concentration model are shown in Figure 22.

![Figure 22. Estimates of *C. jejuni* in Untreated Graywater](image)

The distributions in Figure 22 show that using different indicators to predict the concentration of *C. jejuni* resulted in many different cumulative probability distributions. It is unclear which indicator most accurately estimates the pathogen concentration, indicating the limitations of using fecal indicators to predict pathogen concentrations. The breaks in the distributions shown in Figure 22 are the result of using a lognormal scale to plot the estimated concentration of *C. jejuni*.

The model estimates the median concentration of *C. jejuni* to range from 0.41 to 2100 cells/mL and the estimated mean ranged from 0.47 to 2900 cells/mL. Using coprostanol to estimate the concentration of *C. jejuni* resulted in the lowest estimates while using fecal coliforms resulted in the highest estimates. Using fecal coliforms to estimate pathogen concentration resulted in an estimate of the mean that was approximately 6000 times larger than the estimated mean using coprostanol. It is unclear which indicator would most accurately estimate the pathogen concentration, but it is clear that using different fecal indicators to estimate the concentration of *C. jejuni* can result in very different results. This further illustrates the severe limitations of using fecal indicators to predict pathogen concentrations.

Using the *C. jejuni* concentration distributions as estimated by the pathogen concentration model, the dose and then the risk of infection were then calculated. The dose was based on four exposure scenarios. The risk of infection for the four exposure scenarios is shown in Figures 23 through 26.
The distributions shown in Figure 23 represent the risk of infection from exposure to \textit{C. jejuni} during graywater reuse Scenario 1, which represents accidental ingestion of graywater.

Each of the distributions shown in Figure 23 represent the estimation of the risk of infection from the accidental ingestion of graywater. The median risk of infection ranges from 0.26 to 0.77 infections per year and the mean ranges from 0.26 to 0.78 infections per year as estimated by the model. An acceptable level of microbiological risk in water systems is often cited as less than 1 infection per 10,000 people per year (10^{-4} infections per year) (Fewtrell and Bartram, 2001). Both the median and the mean estimates of risks for Scenario 1 exceed acceptable levels of risk.

The distributions shown in Figure 24 represent the risk of infection from exposure to \textit{C. jejuni} during graywater reuse Scenario 2. Scenario 2 represents indirect ingestion from touching plants and lawns irrigated with untreated graywater.
Figure 24. Risk of Infection from Exposure to *C. jejuni*. Exposure Scenario 2 (Indirect ingestion from touching plants and lawns)

Indirect ingestion of graywater results in multiple distribution curves representing the estimated risks of infection. The median values of risk of infection range from 0.25 to 0.77 per year and the mean estimates range from 0.25 to 0.77 per year. Both the median and the mean estimates of risks for Scenario 2 exceed acceptable levels of risk.

The distributions shown in Figure 25 represent the risk of infection from exposure to *C. jejuni* during graywater reuse Scenario 3, which represents ingestion of graywater from sprays used for irrigation.

Figure 25. Risk of Infection from Exposure to *C. jejuni*. Exposure Scenario 3 (Ingestion of graywater from sprays used for irrigation and washing)
Ingestion of graywater from sprays used for irrigation and washing results in multiple distribution curves representing the estimated risks of infection. For exposure Scenario 3 the median values of risk of infection range from 0.19 to 0.69 per year and the mean estimates range from 0.07 to 0.68 per year. Both the median and the mean estimates of risks for Scenario 3 exceed acceptable levels of risk.

The distributions shown in Figure 26 represent the risk of infection from exposure to *C. jejuni* during graywater reuse Scenario 4, which represents ingestion of crops irrigated with untreated graywater.

![Figure 26. Risk of Infection from Exposure to C. jejuni. Exposure Scenario 4 (Ingestion of crops irrigated by graywater)](image)

Ingestion of crops irrigated with graywater resulted in multiple distribution curves representing the estimated risks of infection. For exposure Scenario 4 the median values of risk of infection range from 0.55 to 0.87 per year and the mean estimates range from 0.55 to 0.86 per year. Both the median and the mean estimates of risks for Scenario 3 exceed acceptable levels of risk.

Using coprostanol to estimate the risks of infection from exposure to *C. jejuni* in untreated graywater resulted in the lowest estimates for each exposure scenario. Additionally, using fecal coliforms to estimate the risks resulted in the highest estimates for each exposure scenario. The range of estimates is largest for Scenario 3. Scenario 3 also represents the exposure scenario with the lowest dose of graywater. This suggests that the risk model is more sensitive to the pathogen concentration for exposure scenarios where low doses of graywater are consumed.

The results of the preliminary risk assessment suggest that the human health risks from exposure to untreated graywater during different reuse scenarios are unacceptably high. This model also suggests that the exposure scenario that presents the highest human health risks is the ingestion of crops irrigated by graywater. These estimates of risk are limited due to uncertainties in the pathogen concentration model.

36
The inputs that contributed the most to the uncertainty of the risk assessment model were evaluated using an importance analysis. The distributions that were evaluated are summarized in Table 12 and the results from this analysis are shown in Figure 27.

### Table 12. Summary of distributions analyzed for importance

<table>
<thead>
<tr>
<th>Importance Analysis</th>
<th>Exposure Scenario</th>
<th>Fecal Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>1</td>
<td>Total Coliform</td>
</tr>
<tr>
<td>#2</td>
<td>2</td>
<td><em>E. coli</em></td>
</tr>
<tr>
<td>#3</td>
<td>3</td>
<td><em>Enterococci</em></td>
</tr>
<tr>
<td>#4</td>
<td>4</td>
<td>Cholesterol</td>
</tr>
</tbody>
</table>

Figure 27 shows the importance of each of the parameters that were input into the risk assessment model and includes intermediate parameters. The inputs with the highest importance are fecal indicators, fecal load, and pathogen concentration, which are all a part of the pathogen concentration model. In addition to importance analysis that was performed for all of the inputs, an importance analysis was performed for the pathogen concentration model.
Figure 28. Dose Model Importance Analysis

Pathogen concentration contributed the most to the uncertainty in the dose model. The importance analysis suggests that the parameters used to estimate the pathogen concentration contribute the most to the uncertainty in the risk assessment model. This indicates that reducing uncertainty in the concentration of pathogens in graywater will greatly reduce the uncertainty in the estimates of risk.

3.5. Updated Risk Assessment

The risks of infection from exposure were estimated using the exposure assessment model and the dose response curve explained in the Materials and Methods section. Although the presence of A. hydrophila was assessed, and A. hydrophila is considered an emerging pathogen in water, the risk of infection was not evaluated because the direct correlation between the dose of A. hydrophila and infection is not clear (Morgan et al. 1995). Thus the risk of infection from exposure to A. hydrophila in untreated graywater could not be estimated.

There were no samples that tested positive for C. jejuni and there was only one sample that tested positive for L. pneumophila. This suggests that the risks of infection from exposure to C. jejuni and L. pneumophila in untreated graywater are low. However, to accurately estimate the risks it was necessary to represent the data from the qPCR analysis using a probability distribution curve.

The presence of C. jejuni and A. hydrophila was below detectable limits for most of the collected samples making it difficult to fit a probability distribution to the qPCR data, thus it was necessary to identify a probability distribution that conceptually matched the characteristics of a microbial environment. A Poisson distribution (Equation 8) was used to represent the data for both C. jejuni and A. hydrophila because it conceptually represents the occurrence of microorganisms in the environment. A Poisson distribution is a discrete distribution in which
any random integer greater than or equal to zero is possible. A discrete distribution was
beneficial in this estimate because a pathogen concentration of zero is possible and could not be
represented accurately by a continuous distribution. (Haas et al, 1999).

\textbf{Equation 8}

\[ P(X) = \frac{\lambda^X}{X!} e^{-\lambda} \]

The Poisson distribution has only one parameter, lambda (\(\lambda\)), which is represented by the mean
pathogen concentration of the graywater samples. To fit Poisson distributions to the
concentration of \textit{C. jejuni} and \textit{L. pneumophila} it was necessary to determine the mean pathogen
concentration of samples. Using qPCR to quantify the pathogen concentration results in a lower
limit of detection determined by the standard curves and was assumed to be the lowest
concentration of targeted pathogen that was detected. Many samples were below the detection
limit for the targeted pathogens and it was assumed that all concentrations between zero and the
detection limit were equally probable. The lower limit of detection was 1.49 cells per qPCR
reaction and 7.29 cells (2.5 \times 10^{-5} \text{ ng}) per qPCR reaction for \textit{C. jejuni} and \textit{L. pneumophila}
respectively. The lower limits of detection for the pathogen concentration (in cells per mL)
needed to be determined using the conversions shown in Equation 9.

\textbf{Equation 9}

\[ \text{Cells per mL} = \frac{\text{Limit of detection (Cells/rxn)} \times \left( \frac{\text{rxn}}{5\muL} \right) \times (100\muL \text{ of Extracted DNA})}{\text{Volume Filtered (mL)}} \]

The volume of graywater filtered varied, resulting in varying lower limits of detection for the
pathogen concentration. The mean lower limits of detection for the pathogen concentration were
0.19 cells per mL and 0.95 cells per mL for \textit{C. jejuni} and \textit{L. pneumophila} respectively. To
determine the mean pathogen concentration of the samples below the detection limit it was
assumed that the concentration could be represented by a uniform distribution ranging from zero
to the detection limit. Thus the mean pathogen concentration of the samples below the detection
limits was half of the mean lower limit of detection for the pathogen concentration. The mean
pathogen concentration was then determined to be 0.097 cells per mL and 0.89 cells per mL for
\textit{C. jejuni} for \textit{L. pneumophila} respectively. The mean concentration of \textit{C. jejuni} was half of the
lower limit of detection for \textit{C. jejuni} because there were no samples above the detection limit
while the mean concentration of \textit{L. pneumophila} was greater than half of the detection limit for
\textit{L. pneumophila} because there was one sample that had a concentration of \textit{L. pneumophila} above
the detection limit. The pathogen concentration was represented by a Poisson distribution with
lambda equal to 0.097 cells per mL and 0.89 cells per mL for \textit{C. jejuni} and \textit{L. pneumophila}
respectively.

For the assessment of the risk of infection from exposure to \textit{C. jejuni}, the mean was very small
because the lower level of detection was small and all of the samples were below the detectable
limit. The resulting risks of infection from exposure to \textit{C. jejuni} are shown by the distributions
in Figure 29. Each of the distributions represents the risk of infection from exposure to *C. jejuni* in untreated graywater from four different exposure scenarios.

![Figure 29. Risk of infection from exposure to *C. jejuni* (Concentration represented using a Poisson distribution fitted to qPCR data)](image)

The resulting risks (Figure 29) are consistent with qPCR data suggesting that for most of the population the risk of infection from exposure to *C. jejuni* in untreated graywater is expected to be zero. However, for a small portion of the population, the risks exceed acceptable limits (1 infection per 10,000 exposures (Fewtrell and Bartram, 2001)). This indicates that pathogen contamination is a valid concern because even small doses can lead to infection. Thus it is suggested that proper barriers be put in place to reduce human exposure to *C. jejuni* in untreated graywater.

The mean concentration of *L. pneumophila* was higher than *C. jejuni* because the lower limit of detection was higher and one of the samples had a concentration of *L. pneumophila* that was above the detectable limit. The risks of infection from exposure to *L. pneumophila* are shown in Table 13.
Table 13. Risk of infection from exposure to *L.*pneumophila (Concentration represented using a Poisson distribution fitted to qPCR data)

<table>
<thead>
<tr>
<th>Exposure Scenario</th>
<th>Estimate of Risk (Year⁻¹)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>5th Percentile</td>
<td>95th Percentile</td>
</tr>
<tr>
<td>#1</td>
<td>0.59</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>#2</td>
<td>0.58</td>
<td>0.99</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>#3</td>
<td>0.36</td>
<td>0.41</td>
<td>0.00</td>
<td>0.88</td>
</tr>
<tr>
<td>#4</td>
<td>0.59</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The estimated risks of infection from exposure to *L.* pneumophila are unacceptably high for the majority of the population. This suggests that proper barriers should be put in place to reduce human exposure to pathogens in graywater.

Additionally the risk of infection from exposure to *L.* pneumophila in untreated graywater was estimated assuming that the sampling from 28 households represented an empirical distribution of the expected concentration of *L.* pneumophila in untreated graywater. For this scenario the distribution was input as discrete values representing the mean value for each of the eighty samples that were evaluated. This suggested a distribution in which there is a one in eighty chance that the graywater would be contaminated by *L.* pneumophila, and if contaminated it was assumed that the value would be the mean value of the triplicate qPCR reactions performed on the graywater sample that tested positive for *L.* pneumophila.

Assuming that the sampling from 28 households represents a distribution of the expected concentration of *L.* pneumophila in untreated graywater, the risk of infection from exposure to *L.* pneumophila was lower than the previous estimate that used a Poisson distribution to represent the concentration of *L.* pneumophila. Using the results from the molecular analysis it was determined that there was a one in eighty chance that graywater would be contaminated with *L.* pneumophila. The resulting risks were then estimated and the median risk of infection was zero for each exposure scenario. The 95th percentile estimate of risk of infection was zero for each exposure scenario. The 99th percentile for the risk of infection was 1 for each exposure scenario. The results suggest that there is approximately a 99 percent probability that the risk of infection from exposure to *L.* pneumophila would be zero. However, it also suggests that 1 percent of the population reusing graywater is expected to be infected by *L.* pneumophila. This indicates that to protect the small portion of the population from infection, proper barriers should be put in place to reduce human exposure to pathogens in graywater.

Each of the estimates provides insight into the human health risks associated with graywater reuse. However there remain uncertainties in all of the assessments of risk that were performed. The risk assessments that relied on a Poisson distribution to represent the concentration of the pathogens are limited because it is difficult to fit a probability distribution to data with many values below the detectable limits. However, using an empirical distribution is limited because it assumes that the 80 samples accurately represent the distribution of the pathogens in graywater and suggests that there are no other possibilities. Additionally, using a Poisson distribution and an empirical distribution to estimate the risks are both limited because the majority of the
samples analyzed by qPCR were below detectable limits for the targeted pathogens. Using an empirical distribution assumes that all samples with a pathogen concentration below the detection limit do not contain any of the targeted pathogen which potentially underestimates the pathogen concentration of the sample. To fit the data to a Poisson distribution required that an assumption be made to determine the pathogen concentration for the samples below the detection limit. It was assumed that all concentrations between zero and the detection limit were equally probable which potentially overestimates the pathogen concentration.

The uncertainty in the model could be reduced by increasing the number of samples or lowering the detection limit. Increasing the number of samples would provide more data points that would provide a better indication of the expected pathogen concentrations. The detection limit is dependent on the available technology and the volume of graywater that is filtered. Increasing the volume of graywater that is filtered would decrease the lower limits of detection of the pathogen concentration providing a more accurate representation of samples with lower pathogen concentrations.

3.6. Microbial Community Analysis and Effect of Storage

3.6.1. Metagenomic Analysis

There were a total of 241,937 pyrosequencing reads, with an average of 1,496 reads per sample, used in the metagenomic analyses. Pyrosequencing reads were used to determine the distribution of members of the Bacteria domain present in the graywater samples. The relative abundance of different taxonomic classes was identified for each of the three sources of graywater. For each source the taxonomic classes were reported as the total percent per source and the average percent per household. The results of this analysis are shown in Figure 30.
Based on this analysis, Alphaproteobacteria, Bacilli, Betaproteobacteria, Bacteroidia, and Negativicutes are the classes of Bacteria that are dominant in graywater generated from bathroom sinks. Actinobacteria, Alphaproteobacteria, Bacilli, and Gammaproteobacteria are the most abundant classes of Bacteria in graywater generated from the laundry. Actinobacteria, Alphaproteobacteria, Bacilli and Gammaproteobacteria represented the majority of Bacteria in graywater generated from the shower/bath.
Using non-metric multidimensional scaling (nMDS) an ordination plot was created to illustrate the similarity between the different samples that were collected. The similarity was determined based on the Bray-Curtis similarity coefficients and was calculated from the presence of operating taxonomical units (OTUs). The results of this analysis are shown in Figure 31.

![Non-metric multidimensional scaling (nMDS) ordination plot](image)

The analysis was based on Bray-Curtis similarity coefficients calculated from the relative distribution of OTUs.

**Figure 31. Non-metric multidimensional scaling (nMDS) ordination plot representing the potential relationship between samples from three graywater sources.**

The samples that are plotted close together (Figure 31) represent samples that are similar, while samples that are plotted far apart represent dissimilarities regarding the OTUs present in the sample. There is a cluster of samples collected from the bathroom sink that indicates that graywater from bathroom sinks are similar. Although there is not one distinct cluster for the laundry and shower/bath samples it appears that most of the samples are near, or similar, to other samples from the same source. Likewise, there appears to be laundry and shower/bath samples that are quite different.

The analysis of similarity (ANOSIM) was conducted on the Bray-Curtis similarity matrix that was developed for nMDS ordination plot. Global and pairwise test were conducted to evaluate the effect of the different graywater sources on the metagenomic structure (Table 14).
Table 11 Results of one-way ANOSIM test based on Bray-Curtis similarity matrix derived from the distribution of microbial communities

<table>
<thead>
<tr>
<th></th>
<th>R Statistic</th>
<th>P-Value</th>
<th>Permutation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global Tests</strong></td>
<td>0.483</td>
<td>0.001</td>
<td>999</td>
</tr>
<tr>
<td><strong>Pairwise Tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bath Sink vs Laundry</td>
<td>0.739</td>
<td>0.001</td>
<td>999</td>
</tr>
<tr>
<td>Bath Sink vs Shower/Bath</td>
<td>0.488</td>
<td>0.001</td>
<td>999</td>
</tr>
<tr>
<td>Laundry vs Shower/Bath</td>
<td>0.182</td>
<td>0.001</td>
<td>999</td>
</tr>
</tbody>
</table>

Significance was set at α=0.05

The results shown in Table 14 indicate a significant difference between the different sources of graywater (P < 0.05). This indicates that the microbial communities are significantly different for each source of graywater. This suggests that the risks of reusing the different sources of graywater are not necessarily the same and thus future evaluations should continue to evaluate each source of graywater separately.

There are many factors that could contribute to the variation between the different sources of graywater. These factors could include the route of contamination, water source, hygiene and laundry products that were used, and the presence of children or pets among others. To evaluate the similarities and differences between the different graywater samples a more extensive survey evaluating various factors should be performed and additional samples should be collected to increase the sample size.

3.6.2. **Coliform Tests**

Total and fecal coliform counts for the samples stored at different times show an increase over time (Figures 32 and 33).
Figure 32. Total Coliforms in stored graywater

The initial concentration of total coliforms present in the stored graywater was below detection (Figure 31). However over time there was an increase in the concentration indicating bacterial growth. This was consistent with previous studies suggesting microbial growth in stored graywater (Rose et al., 1991). Total coliforms were not detected in the graywater from the bathroom sink or the laundry graywater.
The results of the fecal coliform assay (Figure 33) were consistent with the total coliform assays and previously published reports (Rose et al., 1991). The initial concentration of the fecal coliforms present in the stored graywater was very low. However over time there was an increase in the concentration indicating bacterial growth. Fecal coliforms were not detected in the graywater from the bathroom sink water or the laundry.

3.6.3. qPCR analysis

Quantitative PCR targeting the four different organisms was performed for each of the samples that were collected from the storage experiment. Of the samples that were collected none tested positive for \textit{A. hydrophila}, \textit{C. jejuni}, \textit{L. pneumophila}, or \textit{Bacteroides}. The concentration of the general bacteria was also determined using qPCR. The concentrations of the general bacteria are shown in Figures 34 through 37. Each of these figures represents the different types of graywater analyzed in the storage experiment.
Figure 34. Concentration of bacteria in bathroom sink water stored over time

Figure 35. Concentration of bacteria in laundry water stored over time
Figure 36. Concentration of bacteria in shower/bath water stored over time

Figure 37. Concentration of bacteria in combined graywater stored over time

The concentration of bacteria in the different types of graywater resulted in trends that indicate bacterial growth within the first 4 days. Each of the different types of water had an initial concentration of bacteria that ranged from a mean of $5.21 \times 10^4$ to $4.81 \times 10^6$ cells per mL and then resulted in an increase in bacteria in the first 4 days. After the initial growth the bacterial growth stopped and concentrations remained relatively stable. These results are consistent with the coliform assays where there was initial increase in coliform forming bacteria and then the concentrations of bacteria began to stabilize. Graywater typically has a high concentration of easily degraded organic matter that is favorable to bacterial growth (Ottoson and Stenstrom,
2003). However, in the storage experiment there was a finite amount of organic matter. After the organic matter was consumed by the bacteria in the stored graywater, the growth began to slow. It is expected that over time the concentration would begin to decline due to the lack of food available for the bacteria.
4. Conclusions

Graywater is a valuable resource that can be utilized to reduce the amount of potable water needed in the home. However, reusing graywater presents human health risks that must be properly managed. The human health risks were evaluated in this report by quantifying the microbial characteristics of graywater and then assessing the risks of reusing graywater.

Quantitative PCR was used to quantify the concentration of pathogens in untreated graywater. Very few samples tested positive for the targeted pathogens. No graywater samples tested positive for *C. jejuni*, one out of eight tested positive for *L. pneumophila*, and two out of eighty tested positive for *A. hydrophila*. Since pathogens were found in very few graywater samples, if graywater were to be separated by household, pathogen contamination would be infrequent. However, the concentration of the pathogens represented only a small portion of the bacteria present in the graywater, and there is still the possibility that other pathogenic organisms are present in graywater. Additionally it needs to be acknowledged that this study quantified pathogens based on the DNA present in graywater and did not account for the viability of the pathogens present. This method provides a conservative estimate because it is not known if all of the pathogens present in the samples are active.

A preliminary risk assessment showed that using fecal indicators to estimate the concentration of pathogens in graywater introduces uncertainty into the risk assessment model. Quantifying pathogens in graywater using qPCR reduces the uncertainty in the model because it eliminates the need to estimate the pathogen concentration using fecal indicators. The results of the qPCR were used to estimate the human health risks associated with reusing untreated graywater. The estimated risk of infection resulted in estimates that varied significantly from the estimates in the preliminary risk assessment. For a portion of the population, the risk of infection from exposure to *C. jejuni* and *L. pneumophila* exceeded acceptable limits, suggesting that treatment of graywater is needed, and that proper barriers be put in place to reduce human exposure to pathogens in untreated graywater. However, several assumptions had to be made to quantify the risks; these assumptions need to be re-evaluated in future studies using more samples and possibly qPCR with lower limits of detection.

This study was the first metagenomic analysis of graywater using next generation sequencing techniques. In this analysis the microbial communities were identified at the taxonomic class level. The different sources of graywater were determined to be significantly different based on the operating taxonomic units present in collected samples. This suggests that the microbial communities are significantly different for each source of graywater indicating the human health risks associated with reusing graywater could vary from source to source. Thus, human health risks associated with graywater reuse should be evaluated separately for each source of graywater.

The storage experiment was consistent with previous reports (Rose et al., 1991) indicating that stored graywater can provide an environment that is suitable for bacterial growth. In the storage experiment there was an initial increase in the concentration of bacteria in the stored graywater. The concentration of bacteria then began to stabilize likely due to the finite amount of organic matter in the graywater. However, for most graywater reuse scenarios a constant supply of
graywater will pass through storage containers. This suggests that there will also be a continuous supply of organic matter for the bacteria present in the stored graywater. Thus it is recommended that graywater water not be stored for more than 24 hours.

5. Recommendations for Future Studies

This project provided valuable information about the human health risks associated with graywater reuse. However, there remains many important questions that need to be answered. There remains uncertainty in the estimates of risk because it is difficult to fit a distribution to the qPCR data in which the majority of the samples were below detectable limits. This uncertainty can be reduced by collecting more graywater samples and filtering larger volumes of graywater for qPCR analysis.

The risks that were quantified represent the risks of infection from a few targeted organisms. The targeted organisms represented only a small portion of the bacteria in the system. This suggests that there is still a need for additional information regarding the microbial characteristics of graywater. This project provided the first genomic survey of graywater. Metagenomic analysis identified different taxonomic classes present in graywater. Future research should continue to utilize next generation sequencing techniques to identify potential pathogenic organisms present in graywater. We are continuing this work (beyond the original scope of the project) using Illumina sequencing with our collaborators from the US EPA.

While used extensively, DNA-based methods do not determine the viability or the activity of the pathogens present. This affects the risk assessment because pathogens that are no longer active pose little or no human health risks. Future work should include analysis of the RNA present in graywater samples to provide insight into the viability of the organisms present. Variations of PCR that can eliminate extracellular DNA (i.e., DNA that is not associated with intact cells) should also be explored for use in graywater research.

Finally, various graywater treatment methods should be tested, and molecular techniques that take into account viability (such as RNA-based methods, PCR variations, and flow cytometry) should be employed to determine the effectiveness of these treatment approaches.
REFERENCES


APPENDIX

MS Thesis


Conference Papers and Posters