



## **CPS 2011 RFP FINAL PROJECT REPORT**

### **Project Title**

Assessment of *E. coli* as an indicator of microbial quality or irrigation water use for produce

### **Project Period**

January 1, 2012 – December 31, 2012

### **Principal Investigator**

Channah M. Rock, Ph.D.  
University of Arizona, Maricopa Agricultural Center  
channah@cals.arizona.edu, (520) 381-2258

### **Co-Principal Investigator(s):**

Charles P. Gerba, Ph.D., University of Arizona, (520) 621-6906, gerba@ag.arizona.edu

Kelly R. Bright, Ph.D., University of Arizona, (520) 626-8094, bright@u.arizona.edu

### **Objectives**

Phase 1:

1. Determine the best method (most reliable, ease of use, low false positive rate) for *E. coli* detection in irrigation waters based on the comparison of three methods currently available for the detection of *E. coli* in irrigation waters.
2. Determine influence of temperature and salinity (and other environmental factors) on false positive rates of these three methods for accurate *E. coli* detection in irrigation waters.

Phase 2:

1. Develop an exposure scenario (model) for *E. coli* in irrigation waters taking into consideration the type of irrigation method, the irrigated crop, the transfer rate of *E. coli* to the crop, and the *E. coli* survival post irrigation.
2. Estimate the risk of illness from ingestion of various levels of *E. coli* from the proposed irrigation scenarios.
3. Develop a simple, user friendly guideline (program or graph) for estimating risk of infection from the different irrigation scenarios (e.g., different levels of *E. coli* deposited, different crops irrigated). These guidelines will be compared to risks associated with the current guideline of 126 CFU/100 mL.

**Funding for this project provided by the Center for Produce Safety through:**

CDFA SCBGP grant #SCB11069

**Abstract.** The goals of this project were to assess currently used methods for the detection of *Escherichia coli* in irrigation waters used in Arizona and Southern California, and to provide guidelines for a revised *E.coli* standard for irrigation waters used for produce. Currently, there is concern that the false positive rate of *E.coli* detection may be high in these waters giving false indications of the level of risk from enteric pathogens. This may result in unnecessary costly interventions (e.g. disinfection of the water, attempts to limit wildlife access, etc) as well as inaccurate perception of risk among consumers. For this reason it is essential to determine the rate of false positive detection of *E.coli* in waters used for produce irrigation. *E.coli* detection methods were originally developed for assessment of treated drinking water quality and not surface/irrigation waters. Recent research by our group and others has indicated that high temperatures and elevated salinity may result in false positives rates as high as 40% in Arizona and similar climates. The first objective of this project was accomplished by evaluating three commercially available methods for *E.coli* detection to test irrigation waters from three agricultural areas (Yuma and Maricopa, AZ and Imperial Valley, CA) and assessing false positive rates utilizing Polymerase Chain Reaction (PCR) and sequencing of the bacterial isolates. This unique study helped to determine the usefulness of current detection methods for the accurate assessment of *E.coli* contamination in irrigation waters and provides guidance for interpretation of results.

The assessment and confirmation work will, however, have little value without applying these data to the currently proposed *E.coli* guidelines used by the produce industry. Currently, no microbial indicator standards exist for irrigation waters used for produce production in the United States. It has been suggested by the produce industry that the bathing water standard guideline (126 *E.coli* per 100 ml) established by the United States Environmental Protection Agency (USEPA) be used. This guideline was developed from epidemiological studies of bather exposure in recreational waters and has no direct relationship to risk associated with infection or illness rates that might result from produce irrigation waters. Therefore, as a secondary objective to evaluating *E.coli* as a reliable indicator, our team worked to develop a Quantitative Microbial Risk Assessment (QMRA) to assess the risk from consumption of leafy greens following irrigation with waters containing various levels of *E.coli*. The QMRA considers method of irrigation, irrigation timing from harvest time, and other environmental factors that may influence indicator organism or pathogen. This was accomplished using water quality data collected in the first phase of this project coupled with existing information found in the scientific literature. This effort has resulted in a suggested *E.coli* guideline for irrigation waters which reflects irrigation and harvesting practices and is based on human health risk.

***Ultimately this work offers recommendations towards the most reliable methods to be used by the produce industry to assess irrigation water contamination as well as a scientific risk-based E.coli guideline that growers can use to protect public health.***

**Background.** Microbial indicators of water quality have been used for more than 100 years to detect fecal contamination of drinking water and the risk from waterborne

pathogens. Coliform bacteria are still used for this purpose; but coliform bacteria do not always indicate fecal contamination and have potential to re-grow in the environment. The fecal bacterium *Escherichia coli* and enterococci are more likely to be associated with fecal contamination and a correlation between bathing in recreational waters and illness has been established in several epidemiological studies.

Various methods have been developed to detect fecal bacteria in water. Many of these methods specifically utilize an enzyme-based chromogenic media for the detection of *E.coli* in water. Though molecular analyses have been shown to be more sensitive and specific for detection of target microorganisms (Fratamico and Bagi, 2001), enzyme-based media allow for better interpretation of the analysis results, are easier for water utilities to use, and have the advantage of providing response in a shorter time compared with other methods (Edberg and Kontnick, 1986; Sartory and Howard, 1992). Since 2002, the USEPA has approved ten enzyme based *E.coli* detection tests for examination of drinking water (Olstadt et al., 2007), most of which are based on the detection of the *E.coli*-associated enzyme  $\beta$ -D glucuronidase. Despite their widespread use, the application of enzyme-based media for detection of *E.coli* in complex environmental matrices, such as surface/irrigation waters, can present challenges, as the majority of these methods were developed for the detection of *E.coli* in relatively clean drinking or surface water samples in temperate climates (Eccles et al., 2004). While false positive (organisms identified as *E.coli* that actually are not) and false negative (actual *E.coli* that are mis-identified by the method) results are rare in clean treated drinking water, (Edberg et al. 1988) recent research has indicated that they may be much greater in wastewater and surface waters (Chao, 2006; McLain and Williams, 2008).

Utilizing an enzyme-based selective medium, McLain and Williams (2008) reported that false positive rates in *E.coli* identification in treated wastewater in Arizona within a constructed wetland showed sharp seasonal differences, averaging 35% during warm summer months but rising sharply to 75% in the late fall and winter. This finding of high false positive rates during some seasons, and the widespread use of enzyme-based media by laboratories conducting water quality analyses, potentially results in water managers relying on data which does not accurately represent the actual water quality. In a more recent study McLain et al. (2011) found false positive rates for *E.coli* ranging from 4.0% to as high as 48.8% in Arizona waters and soil. Higher false positive rates were correlated with cooler air temperatures and lower salinity levels. Chao (2006) found false positive rates for *E.coli* as high as 36% in tropical surface waters.

In addition to environmental factors that may inference false positive rates there is a wide variability in the various methods used for *E.coli* detection (Olstadt, J. et al. 2007) and their ability to suppress false positive results.

Exacerbating these concerns over method variability and false positive rates, recent outbreaks of *E.coli* O157:H7 and *Salmonella* in fresh produce, coupled with heightened media coverage of these events, have thrust fruit and vegetable safety into the forefront of public attention. Several produce outbreaks have been known or suspected to have

arisen from contamination where irrigation water was suspect as the source (Gerba and Choi, 2009).

In the winter of 2010, romaine lettuce contaminated with *E.coli* O145 caused a multistate outbreak of enteric disease. The contaminated produce was traced back to a single processing facility in Yuma, Arizona. The resulting report from the US Food and Drug Administration (US FDA) hypothesized that a septic system from a trailer park had leaked human waste into an irrigation canal, and this was the presumptive source of the disease-causing *E.coli* (US FDA, 2011). The US FDA report contained several suggested preventative control strategies, among them the “*development and implementation of microbiological monitoring protocols for the canal water system.*” Arizona and California growers have been affected by a decline in consumption of fresh produce and are faced with unprecedented scrutiny and a greater need to implement on-farm food safety assurance systems.

As stated earlier, there currently is no microbial indicator standard(s) for irrigation waters (other than reclaimed municipal wastewater) used for produce production in the United States. Forty years ago, Geldreich and Border (1971) suggested a standard for irrigation waters of 1,000 fecal coliforms per 100 ml based on the likely presence of *Salmonella* in waters when these numbers of fecal coliforms were exceeded. It has been suggested by the produce industry that the bathing water standard guideline of 126 *E.coli* per 100 ml guideline developed by the USEPA be used. This guideline was based on epidemiological studies among bathers on recreational waters and has no direct relationship to risk associated with infection or illness rates that might result from irrigation waters used for produce production. Thus, there is currently no scientific basis for their use in irrigated agriculture. Factors such as irrigation methods and degree of pathogen transfer to the produce and survival of pathogens need to be taken into consideration.

Quantitative Microbial Risk Assessment (QMRA) is a process that allows an assessment in a quantitative fashion for estimation of the risk of infection and illness from a pathogenic microorganism. It has seen widespread application in water and food (Regli et al., 1991; Rose et al, 1995; Jaykus, 1996; Haas et al., 1999; Stine et al., 2005a). QMRA involves four basic steps: (1) hazard identification (the pathogen in question), (2) dose-response (the probability of infection from ingestion of a given number of organisms), (3) exposure assessment (the number of organisms ingested), and (4) risk characterization (estimation of the probability on infection or illness). Our team has previously used QMRA to estimate the probability of infection from hepatitis A virus and *Salmonella* in irrigation waters used to irrigate cantaloupe and iceberg lettuce by various irrigation methods (Stine et al, 2005a). QMRA has also been used to estimate risks and develop guidelines for the use of reclaimed wastewater for food crop irrigation (Pettersen et al., 2001) and has been recommend by the World Health Organization for developing treatment guidelines (WHO, 2006).

QMRA is usually used to estimate the risk from the ingestion of a known pathogen by water or food. However, if the level of exposure to an indicator (*E.coli*) is known and risk

of illness is known, the risk of illness can be estimated. This is the case with the *E. coli* standard for bathing water where epidemiological studies have shown that when the *E. coli* level in the bathing water is 126 *E. coli* per 100ml, the risk of gastrointestinal illness is 1 in 50 (Cabelli, 1989). By knowing the ingestion rate of water during bathing (38 ml from studies), the number of *E. coli* ingested can be used to estimate the probability of illness, without knowing which pathogen caused the illness. The risk of illness from exposure of *E. coli* during bathing is known for different levels of *E. coli* in the water, which will allow us to predict risk of infection from ingestion of different levels of *E. coli* deposited on fresh vegetables through contaminated irrigation water.

**Research Methods and Results.** The following research methods and results are divided into two sections representing the two defined phases of the project.

**Research Methods Phase 1:** Three different methods for detection and quantification of *E. coli* in water were evaluated during this study. These methods included: (1) MI agar (Becton-Dickinson, Franklin Lakes, NJ); (2) IDEXX Colilert Quanti-Tray® (IDEXX Laboratories, Westbrook, ME); and (3) m-ColiBlue24® broth (Hach Co., Loveland, CO).

Field Sampling: Samples were collected in three different agricultural areas: Maricopa and Yuma (AZ), and growing regions of Imperial Valley (CA). These three locations represent a significant portion of winter leafy green production in the United States. Sample collection was carried out over the course of one year during the winter growing season and additional select times of year to assess the effects of temperature, salinity and other environmental factors (e. g. sunlight intensity, precipitation) on method variability.

Over the course of the project a total of 150 1L grab samples were collected at each of the 3 locations for a total of 450 samples. Following collection, water samples were placed on ice for transport to the laboratory. Sampling sites were determined based on relative distance up-stream from the irrigation practice and each production field. This was done in collaboration with University of Arizona faculty at the Yuma Agricultural Center (YAC), the Maricopa Agricultural Center (MAC) and cooperating industry partners and irrigation districts.

Laboratory Analysis: In the laboratory, water samples were divided equally for testing by the three above methods for enumeration of *E. coli*. Methods including MI agar and m-ColiBlue24® broth involve sample processing using the membrane filter technique, which is widely accepted and approved as a procedure for monitoring water microbial quality in many countries (Rompré et al., 2002; USEPA, 2002a). Water samples were filtered in three dilutions (100, 10, and 1.0 mL) through Whatman gridded filters (0.45 µm pore size, 152 mm diameter) (Whatman International Ltd., Kent, UK). For MI Agar, filters were placed onto 60 mm Petri plates with prepared MI Agar with cefsulodin added at 5 mg L<sup>-1</sup> 153 to inhibit growth of Gram positive organisms and selected non-coliform Gram-negatives (e.g., *Pseudomonas* spp.). Plates were incubated at 37° C for 24-36 h before counting dark blue colonies presumptive for *E. coli*.

Water samples aliquots for IDEXX Colilert Quanti-Tray® were processed according to manufacturer instructions within 6 hours of sample collection and trays were read after 18 hours incubation at 37° C. Only wells fluorescing “blue” color were selected for false positive confirmation.

For the method involving m-ColiBlue24® broth the filters were placed in a Petri plate containing an absorbent pad soaked in m-ColiBlue24® broth. Petri plate samples were inverted and incubated at 35° C for 24 hrs. Following incubation colonies showing a blue color will be selected as presumptive for *E.coli*. m-ColiBlue24® broth has been approved by the EPA for monitoring drinking water and can also be used to detect coliforms and *E.coli* in other types of water (e.g., bottled, surface, ground, well) (US EPA, 40 CFR Parts 141, 143).

In addition to the water microbiology mentioned above, additional analyses to assess factors contributing to false positive identification were performed. These included heterotrophic plate count (HPC), total alkalinity, salinity, temperature, pH, turbidity, and total dissolved solids. All assays were performed according to Standard Methods (Standard Methods, 1998) at the University of Arizona Microbial Water Quality Laboratory (Maricopa, AZ) or the University of Arizona Environmental Research Laboratory (Tucson, AZ).

Due to the space limitation of this report water quality correlation data is not presented here and will be referenced in the Appendices.

Confirmatory Analysis of *E.coli*: Individual bacterial colonies for confirmatory analysis of *E.coli* and non-*E.coli* were collected from approximately 1/3 of the total number of samples analyzed. Blue (presumptive *E.coli*), colorless (presumptive non-*E.coli*) and red (presumptive coliforms (non-*E.coli*)) from MI Agar and m-ColiBlue24® broth were collected and re-streaked onto TSA (Trypticase soy agar), as were 10 µL aliquots removed from Quanti-Tray® fluorescent (presumptive positive for *E. coli*) and non-fluorescent (negative) wells. The remaining broth was extracted from the selected well (both positive and negative) and added to test tubes containing sterile EC broth (BD, Franklin Lakes, NJ) and incubated at 44.5C for 24 h. Incubation tubes testing positive for *E.coli* (turbidity and gas production) were streaked onto TSA. Individual isolates were frozen at -80° C until re-growth and molecular analysis.

Identification of False Positive and False Negative Organisms: Of the over 900 colonies frozen for sequencing 20% were randomly selected for PCR and DNA Sequencing using a Random Number Generator (<http://stattrek.com/statistics/random-number-generator.aspx>). For identification of bacterial isolates by DNA sequencing, frozen isolates presumptive for *E.coli* and non-*E.coli* were streaked onto TSA plates for collection of a single isolate that was used as a template for Polymerase Chain Reaction (PCR). PCR amplicons were generated using universal primers (63f and 1389R) specific for a 16S region common to all known bacteria. PCR reactions contained 200nM of each primer (Forward: 5'-CAG GCC TAA CAC ATG CAA GTC-3'; Reverse: 5'-ACG GGC GGT GTG TAC AAG-3'), 12.5 µL Thermo Scientific Master Mix,

1.0  $\mu\text{L}$  of bacterial DNA lysate and sterile water to 25  $\mu\text{L}$ . Reactions were run on an Eppendorf Master Cycler-Pro (Eppendorf North America, Happauge, NY) using a program consisting of the initial denaturation at 94°C for 2 min, followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min. Each PCR run included positive and negative control tubes with template *E. coli* (ATCC #25922) and nuclease-free water, respectively. Amplification product was detected by electrophoresis in a 2% agarose gel, stained with SYBR™ Safe DNA Gel Stain (Invitrogen Life Technologies, Carlsbad, CA) and visualized by transillumination by UV light.

The 1326-bp amplicons were then sequenced in both directions using an Applied Biosystems 3730 DNA Analyzer at the University of Arizona Genetics Core, Tucson, AZ. Retrieved sequences were imported into Sequencer 5.1 (Gene Codes Corporation, Ann Arbor, MI) where the ends were trimmed and the forward and reverse primer sequences were made into contiguous sequence FASTA files. These files were then uploaded to RDP-II (Ribosomal Database Project, Michigan State University, Ann Arbor, MI) for species identification.

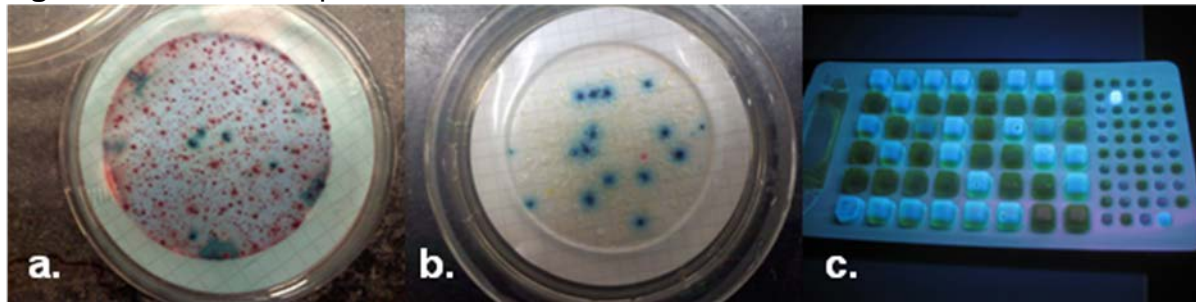
Sequence Analysis: The Ribosomal Database Project (RDP-II) is hosted by the Center for Microbial Ecology at Michigan State University and provides data, programs and services related to the ribosome. The ribosomal RNA sequences in the RDP-II alignments are drawn from major sequence repositories like GenBank and European Bioinformatics Institute and from direct submissions to the RDP. This tool can be found at <http://www.cme.msu.edu/RDP>. For this project the University of Arizona Genetic Core (UAGC) provided files containing the DNA sequences. UAGC performed high volume sequencing on PCR product generated from bacterial cultural field isolates using the oligonucleotide primers designed to anneal at bacterial 16s rRNA genes 63f (5'-CAGGCCTAACACATGCAAGTC-3') and 1389r (5'-ACGGGCGGTGTGTACAAG-3'). These files were then downloaded and imported into Sequencer 5.1 (Gene Codes Corporation, Ann Arbor, MI) where the ends were trimmed and the forward and reverse primer sequences were made into contiguous sequence FASTA files. These files were then uploaded to RDP-II for species identification. RDP-II offers aligned and annotated rRNA sequence data and analysis services to the research community. This database is updated regularly and maintained 2.3 million aligned and annotated quality-controlled rRNA sequences and associated information as of June 2012. The tools for taxonomic classification (RDP Classifier), sequence match and comparison of taxonomic content (Library Compare) were all used to analyze this projects DNA sequences.

Statistical Analysis: All means and standard errors were calculated using un-transformed data. Prior to regression analyses, all microbiological (CFU) data was checked for normality using probability plots, and data not conforming to a normal distribution was log-transformed using the Johnson transformation (Luh and Guo, 2001). Effects of chemical, physical, and biological water quality parameters on appearance of false positives were analyzed by binary logistic regression. Each isolate analyzed by PCR was characterized as 0 = *E.coli* or 1 = false positive, this character will be used as the binary dependent variable in models identifying significant relationships between false positives and each environmental factor (entered into each model as a

continuous variable). A significance level of  $\alpha = 0.05$  will be used for all statistical tests. Statistical analyses will be performed using Minitab 13.32 Statistical Software (Minitab Inc., State College, PA).

**Research Results Phase 1:** Results reveal *E.coli* in irrigation waters in all agricultural areas sampled in this study, including exceedances of the LGMA guideline of 126 *E.coli* per 100 mL. While all three methods have identified *E.coli* in irrigation waters, methods including MI agar, and IDEXX Colilert Quanti-Tray® have shown the most straightforward results for interpretation, while blue colonies on m-ColiBlue24® broth plates are typically not well defined, making it difficult to differentiate between a single colony or multiple colonies, which could over- or underestimate the *E. coli* in the sample (see figure 1 below).

**Figure 1.** Method comparison



(a.)m-ColiBlue24® broth; (b.)MI Agar; and (c.)IDEXX Colilert Quanti-Tray®.

The following figures represent *E. coli* counts in irrigation waters of irrigation canals in the Yuma areas are shown in Figure 1(appendix A), while Figures 2 and 3 (appendix A) show the *E. coli* counts for the Imperial Valley, CA and Maricopa, AZ respectively. Figure 4 (appendix A) shows the *E. coli* counts determined for all sites. The maximum value plotted in Figures 1 through 4 is the maximum count obtained for the three different enumeration methods for each sample. The 126 level is shown as a line passing through the value of 126 *E. coli* count per 100 ml.

**Statistical Analysis:** All statistical and data analysis were conducted using the different packages in the programming Language R (The R Project for Statistical Computing, 2012).

**Testing Significant Differences between the 3 Lab Methods:** To determine if there is statistical significant difference between the measurements obtained using the three different *E. coli* detection methods for all locations, analysis of variance (ANOVA) was conducted using the Language R (The R Project for Statistical Computing, 2012). It is important to note that the actual values for *E. coli* per 100 ml obtained from the different methods were used to conduct the ANOVA. Table 1 shows the different statistics obtained from the different methods used to test for *E. coli*.

The condition of normality of the measured data was tested and the Q-Q plot was developed as shown in Figure 5 (appendix A). Figure 5 shows that the data did not

come from a normally distributed population since most of the data lies outside the dotted lines on the graph representing the 5<sup>th</sup> and 95<sup>th</sup> percentiles. It was determined to transform the *E. coli* counts using the log<sub>10</sub> and repeat the analysis on the transformed data. Table 2 shows the statistics for the *E. coli* counts after applying the log<sub>10</sub> transformation to the data.

**Table 1.** Statistics for the Different lab Methods

Statistics	Method		
	IDEXX Colilert Quanti-Tray®	m-ColiBlue24® broth	MI Agar
Number of Measurements (N)	446	438	439
Arithmetic Mean	37.81	16.18	36.45
Standard Deviation	113.44	37.26	62.82

**Table 2.** Statistics for the Different Methods Based on Log<sub>10</sub> Transformation of Measurements

Statistics	Method		
	IDEXX Colilert Quanti-Tray®	m-ColiBlue24® broth	MI Agar
Number of Measurements (N)	446	438	439
Geometric Mean	12.78	6.97	15.25
Standard Deviation of Log <sub>10</sub> Transformed Data	0.5932	0.5403	0.5931

To determine if the three lab methods have the same variance, we conducted the Bartlett test of homogeneity of variances. The analysis produced a *p*-value of 0.08275 which is greater than 0.05 indicating that there are no significant differences between the variances of the 3 methods.

Finally, analysis of variance methodologies can be sensitive to the presence of outliers. One can test for outliers. The test indicates that there are no Studentized residuals with Bonferonni *p*-value of 0.2715 meaning that there are no outliers measured values among the three methods.

Since the normality test indicates that the log<sub>10</sub> transformed measured lab results for the 3 methods have normal distribution, equal variances and have no outlier values, then analysis of variance can be completed as shown below.

To determine if there is a significant difference in the *E. coli* counts per 100 ml of irrigation waters for the different testing methods, ANOVA was conducted and a *p*-value of  $2 \times 10^{-16}$  was determined indicating that there are significant differences between the different methods. To compare each method to the other 2 methods, multiple comparison ANOVA tests were conducted and Table 3 shows the *p*-values.

**Table 3.** Results of the ANOVA Multiple Comparison Tests\*

Method 1	Method 2	<i>p</i> -value
m-ColiBlue24® broth	IDEXX Colilert Quanti-Tray®	<b>&lt; 0.0001</b>
MI Agar	IDEXX Colilert Quanti-Tray®	0.1160
MI Agar	m-ColiBlue24® broth	<b>&lt; 0.0001</b>

\***Note:** *Italicized and bold p-values indicate statistical significant difference*

The *p*-values in Table 3 show that there are significant differences between *E. coli* counts measured using m-ColiBlue24® broth and those measured using Colilert Quanti-Tray®; and between those measured using Colilert Quanti-Tray® and MI Agar. However, with *p*-value of 0.1160 there are no significant differences between *E. coli* counts measured using MI Agar and Colilert Quanti-Tray®.

These are important findings when recommending the most appropriate method for reliable *E. coli* assessment in irrigation waters. Our results show that while there is little variance in each of the three methods analyzed culturally, that when compared, the MI and the Colilert seem to perform similarly and have no significant difference.

Confirmatory Analysis of *E. coli*: As stated above approximately 20% of the 900 colonies archived from each of the three methods were processed for DNA sequencing. This was done in order to determine if there were differences in the variance between methods molecularly when compared to the cultural results.

The following table (Table 5) describes the sequencing results for each of the three methods evaluated in this study. Results are presented as True False Positives, True Positives, True False Negatives, and True Negatives. To aid in data evaluation, a table (Table 4) has been provided to aid in interpretation of the results.

**Table 4.** Description of Result

Name	Cultural Result	Molecular Result	Actual Result
True False Positive	Presumptive Positive	Negative	Non- <i>E. coli</i>
True Positive	Presumptive Positive	Positive	<i>E. coli</i>
True False Negative	Negative	Positive	<i>E. coli</i>
True Negative	Negative	Negative	Non- <i>E. coli</i>

Results indicate that while each of the three methods do fairly well at identifying *E. coli* in irrigation waters across all three regions, that some of the methods perform better than others and have higher or lower efficiencies. The Colilert Quanti-Tray® performed with the highest rate of accuracy with 49% of the time calling a true positive followed by MI Agar and m-ColiBlue24® broth at 33% and 29% respectively. These results are similar

to past research findings by McLain et al. (2011). However, each of the three methods seemed to have elevated False Positive rates indicating the difficulty in accurately assessing *E.coli* concentrations. This could be due heavily to analyst variability and points towards the need in methods to be straight forward and user friendly. False Positive rates ranged from 53% to 71% with m-ColiBlue24® broth performing the worst. False Positive results have large implications of over-estimating actual concentration of *E.coli* in a water sample. This could lead to costly interventions, such as non-use of water, delayed irrigation, and rejected product just to name a few. Lastly, False Negative results were analyzed for all three methods. While both the MI Agar and the m-ColiBlue24® broth had high efficiencies at calling a true negative colony negative, the Colilert Quanti-Tray® indicated a False Negative rate of 35%. This means that in 35% of the wells that were non-fluorescing (negative) *E.coli* was detected and confirmed my molecular sequencing. According to the manufacturer, as little as one colony per well should induce a color change and result in a positive reading, however this was not observed in our evaluation.

**Table 5.** Method Efficiency

Method	Confirmation Result			
	True Positive <sup>¶</sup>	True False Positive <sup>‡</sup>	True Negative <sup>±</sup>	True False Negative <sup>§</sup>
MI Agar	33%	66%	100%	0%
m-ColiBlue24® broth	29%	71%	100%	0%
IDEXX Colilert Quanti-Tray®	49%	53%	64%	35%

¶True Positive: Positive for *E.coli* on growth media and *E.coli* DNA

‡True False Positive: Positive for *E.coli* on growth media and not DNA

±True Negative: Negative for *E.coli* on growth media and DNA

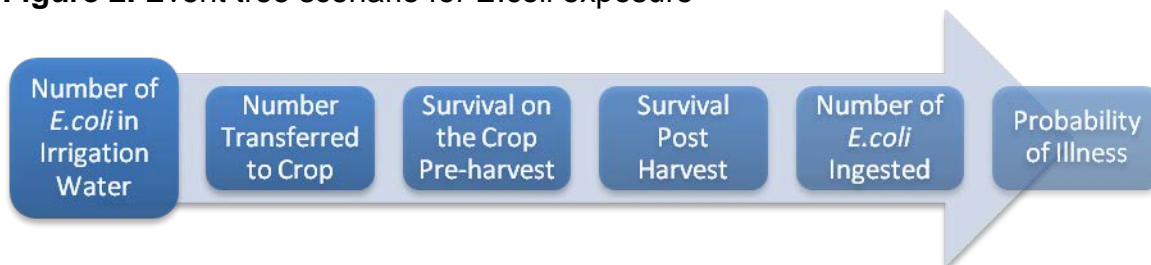
§True False Negative: Negative for *E.coli* on growth media and not DNA

**Research Methods Phase 2:** The second phase of this project focused on the development of guidelines to estimate risk of illness from different levels of *E.coli* in irrigation waters. Using QMRA, our team assessed the risk from consumption of leafy greens following irrigation with waters containing various levels of *E.coli*. Factors considered were method of irrigation, irrigation scheduling in relation to harvest time and other environmental factors that may influence survival of the indicator organism or pathogen.

Development of risk scenarios: The probability of illness from exposure to different concentrations of *E.coli* in recreational waters is known (Cabelli, 1989). The observed illness may be caused by a variety of pathogens including bacteria, viruses and protozoa. However, the relationship between *E.coli* and pathogens causing human illness can be used to estimate probability of illness from different levels of exposure to *E.coli*. A certain amount of uncertainty exists in this approach because the concentrations and types of pathogens in irrigation waters may be different than those

in bathing waters. However, many of the irrigation waters in the Western United States originate from reservoirs and/or rivers where recreation is common. The QMRA we developed is dependent on knowledge of the level of *E. coli* ingested on the produce by the consumer, determined from research reports showing transfer rates of *E. coli* from irrigation to the produce and its survival on the produce. This is dependent on both the type of irrigation method and type of produce. This information was available from the literature (Pettersen et al., 2001) and our own studies (Stine et al. 2005a; 2005b; Stine et al. submitted for publication). We have used a similar approach to estimate risks from produce by *Salmonella* and hepatitis A virus (Stine 2005a). This “event tree” approach has also been used by Gale (2003) to estimate risk from pathogen contamination of food crops (Figure 2). The scenarios for this study were developed for various irrigation delivery systems (e.g., drip, furrow, and sprinkler) and focus on impacts to leafy greens.

**Figure 2.** Event tree scenario for *E. coli* exposure



Estimate the risk of illness from ingestion of various levels of *E. coli* from the irrigation scenarios: Since the yearly consumption of various produce in the United States is known, the yearly risk of illness can be determined (Stine et al., 2005a). This information was modeled for each type of irrigation method, and the information was then incorporated into graphics to allow estimates of risk based on different scenarios.

It should be noted that a limitation to this approach is that the pathogens found in drinking water may be different than those found in bathing waters and that the die-off of the pathogens on the crop pre- and post-harvest may be different than *E. coli*. Thus, the final step in Phase 2 involved an assessment of the plausibility of the QMRA results using *E. coli*. Data used was available from literature and our own studies on the occurrence of enteric pathogens and *E. coli* in a range of water samples (Thurston-Enriquez et al, 2002; Kayed, 2003). Using this data our team estimated the level of risk from the *E. coli* and pathogens in the same water. While, one cannot expect a direct correlation, these results will demonstrate that the relationship between *E. coli* and enteric pathogens can be conservatively estimated within a range in water quality found in irrigation waters.

**Research Results Phase 2: Evaluating Probability of Risk of Infection for Field Measured *E. coli* Concentrations.** In this evaluation, the maximum *E. coli* concentrations measured in the field using the 3 different methods were utilized. The measured concentrations represent the irrigation water concentrations of *E. coli* per 100 ml of irrigation water. Figure 18 (appendix A) represents the annual probability of risk of infection for *E. coli* if a person directly ingested 100 ml of irrigation water on daily basis.

Notice that the three marks represent 77, 126, and 10,000 concentration of *E. coli* per 100 ml of irrigation water (Johnson, 2001; US EPA 1987; WHO 2006). British Columbia, Canada has announced their intention to use a similar approach to the US EPA in setting generic *E. coli* criteria for irrigation water used on produce consumed raw. Their irrigation criteria (less than or equal to 77 CFU per 100 ml geometric mean) are the same as and were derived from those used for primary-contact recreation. Alternatively, based upon a WHO analysis of tolerable risk for irrigation water, the minimum microbial quality for water used on root crops that are eaten raw is 1000 CFU generic *E. coli* per 100 ml (10,000 CFU generic *E. coli* per 100 ml in leaf crops). According to the WHO analysis, using water of this microbial quality is dependent upon a 2 log reduction due to die-off between last irrigation and consumption (includes die-off in the field and during distribution) and a 1 log reduction attributed to washing prior to consumption.

Stine et al. (2005) measured in the field the fraction of *E. coli* that is transferred to lettuce when irrigated with water inoculated with *E. coli*. The maximum ratio for furrow irrigation systems was found to be  $1.1 \times 10^{-4}$  and for subsurface irrigation systems  $8.8 \times 10^{-7}$ . Stine et al. (2011) found that the transfer fraction for sprinkler irrigation systems to be  $1.1 \times 10^{-2}$ . Stine et al. assumed in both the 2005 and 2011 studies that the adjusted per capita consumption of lettuce in the United States is 4,416.5 grams and that the consumption of lettuce occurs one day after irrigation. Multiplying each fraction for the respective irrigation system by the maximum concentration of the *E. coli* measured in the field yields figures 19 through 21 (appendix A) that represent the annual risk of infection from eating an average of 12.1 grams of fresh lettuce per day when the lettuce is irrigated via the specific irrigation system.

The results of risk assessment are indicative that irrigation of lettuce with water containing 126 *E. coli* per 100 ml would result in a risk of less than 1:300,000 from a onetime consumption of the lettuce for all types of irrigation methods. The risks for subsurface drip irrigation were less than  $10^{-9}$ . Even considering multiple exposures (consumption of contaminated lettuce several times) these risks would be exceptionally low. It would appear based on these assumptions, the risk of transmission of *E. coli* 0157:H7 using irrigation water with 126 *E. coli* per 100 ml would be exceptionally low.

Epidemiology Based Risk Assessment. The U.S. EPA (1984) developed a health affects criteria for swimming in fresh recreational waters using *E. coli* as an indicator using epidemiological data on gastrointestinal illness (GI) among bathers. They suggested as a standard 126 *E. coli* per 100ml as the recreational standard. They found that this corresponds to 8 illnesses per 1,000 bathers. The cause of the illness was not known, but believed to be of largely viral origin. Additionally, all of the bathing waters studied where impacted by sewage discharges. Figure 22 (appendix A) adapted from U.S. EPA (1984) shows probability of gastrointestinal (GI) illness as a function of *E. coli* density per 100 ml of swimming water.

Figure 22 (appendix A) shows that 8 in 1000 swimmers will have GI illness if they swim in fresh recreational water that has *E. coli* density of 126 per 100 ml. It is assumed that

bathers swallow 10 ml or ingest 12.6 *E. coli* when they swim in water containing 12.6 *E. coli* and this will result in 8 GI illnesses per 1,000 persons.

Using a proportional linear approach with Figure 22 (appendix A), if irrigation water has *E. coli* density of 126 per 100 ml (or 12.6 *E. coli* per 10 ml), and based on Stine et al. (2005), 0.00011 of the 126 *E. coli* per 100 ml (or 12.6 *E. coli* per 10 ml) will be transferred to lettuce for furrow irrigation system and  $8.8 \times 10^{-7}$  of the 126 will be transferred to lettuce for subsurface drip irrigation system. That corresponds to a risk of GI illness of 1.1 in 100,000 for furrows and 9 in 100,000,000 for subsurface irrigation system. For sprinkler irrigation system and based on Stine et al. (2011), 0.011 of the 126 *E. coli* per 100 ml (or 12.6 *E. coli* per 10 ml) will be transferred to lettuce resulting in a risk of GI illness of 1.1 in 1,000.

**Table 6.** Relative risk of GI illness in lettuce based on irrigation practice\*

<b>Irrigation Practice</b>	<b>Relative Risk</b>
Subsurface Irrigation	9 in 100,000,000
Furrow Irrigation	1.1 in 100,000
Sprinkler Irrigation	1.1 in 1,000

\* Assumes 126 *E. coli* per 100 ml of irrigation water.

Overall, based on these assumptions the risk of illness from consumption of lettuce with irrigation water would be low. This assessment only applies to lettuce. Uncertainty in the estimate could be reduced by better data on transfer efficiency of spray irrigation since the data of Stine et al. (2011) was based on the use of pesticide spray from a container. Also, this assessment is only a onetime event and multiple contamination or multiple days of consuming the same head of lettuce would increase the risk. Still, irrigation water containing 126 *E. coli* per 100 ml for lettuce would appear to present a minimal risk for furrow and subsurface drip. However, further research on contamination of lettuce by spray irrigation appears warranted to reduce uncertainty in the risk estimate.

**Outcomes and Accomplishments.** Overall the research team was able to evaluate three currently used methods for the accurate assessment of *E. coli* in irrigation waters used for produce. Our concluding findings indicate that while all three methods are able to detect *E. coli*, the variance between them is great and that the IDEXX Colilert Quanti-Tray® seems to be the best choice when given an option. However, it is important to note that while the IDEXX Colilert Quanti-Tray® performed well at reducing False Positives rates, that False Negative rates were higher than the two other methods compared. This is important information for the industry and testing labs currently utilizing these methods for *E. coli* assessment.

Additionally, through our comprehensive evaluation, a more robust QMRA analysis was performed using actual *E. coli* data collected throughout the growing region. This is the first study of this kind using actual environmental data and applying it to current regulatory guidelines for irrigation waters used for produce. The results of this risk assessment will be shared industry wide. To aid in information sharing, the final results of this project will be presented at the Center for Produce Safety meeting in Rochester,

NY in June, 2013. Also, our team is currently working to finalize a *Risk Communication Packet* which contains a series “fact sheets” through the University of Arizona Cooperative Extension. This packet summarizes the research results but also includes sections on frequently asked questions, key message points, definitions regarding what are risk assessments, relative risk, and how water quality and irrigation risks compare to other risks commonly observed by the general population.

### Summary of Findings and Recommendations.

- Results reveal *E.coli* in irrigation waters in all agricultural areas sampled, including exceedances of the LGMA guideline of 126 *E.coli* per 100 mL. All three methods have identified *E.coli* in irrigation waters, but methods including MI agar, and IDEXX Colilert Quanti-Tray®, have shown the most straightforward results for interpretation, while blue colonies on m-ColiBlue24® broth plates are typically not well defined, making it difficult to differentiate between a single colony or multiple colonies, which could over- or underestimate the *E. coli* in the sample.
- Our study indicates that there are significant differences between *E. coli* counts measured using m-ColiBlue24® and those measured using IDEXX Colilert Quanti-Tray®; and between those measured using IDEXX Colilert Quanti-Tray® and MI methods. However, there are no significant differences between *E. coli* counts measured using MI Agar and IDEXX Colilert Quanti-Tray®.
- The IDEXX Colilert Quanti-Tray® performed with the highest rate of accuracy with 49% of the time calling a true positive followed by MI Agar and m-ColiBlue24® broth at 33% and 29% respectively.
- Each of the three methods seemed to have elevated False Positive rates indicating the difficulty in accurately assessing *E.coli* concentrations. This could be due heavily to analyst interpretation and points towards the need in methods to be straight forward and user friendly. False positive rates ranged from 53% to 71% with m-ColiBlue24® broth performing the worst.
- According to the QMRA, if irrigation water has *E. coli* density of 126 per 100 ml (or 12.6 *E. coli* per 10 ml), and based on Stine et al. (2005), 0.00011 of the 126 *E. coli* per 100 ml (or 12.6 *E. coli* per 10 ml) will be transferred to lettuce for furrow irrigation system and  $8.8 \times 10^{-7}$  of the 126 will be transferred to lettuce for subsurface drip irrigation system. That corresponds to a risk of GI illness of 1.1 in 100,000 for furrows and 9 in 100,000,000 for subsurface irrigation system.
- For sprinkler irrigation system and based on Stine et al. (2011), 0.011 of the 126 *E. coli* per 100 ml (or 12.6 *E. coli* per 10 ml) will be transferred to lettuce resulting in a risk of GI illness of 1.1 in 1,000.
- Irrigation water containing 126 *E. coli* per 100 ml for lettuce would appear to present a minimal risk for furrow and subsurface drip. However, further research on contamination of lettuce by spray irrigation appears warranted to reduce uncertainty in the risk estimate.

## **APPENDICES**

**Appendix A. Publications and Presentations**

**Appendix B. Budget Summary**

**Appendix C. Tables and Figures** (See references in text to Appendix A.)

**Appendix D. Suggestions to CPS**

## Appendix A.

### 1) List of Publications:

- a) Statistical Methods and Analysis, Quantitative Microbial Risk Assessment (Technical Report, pgs.18-50)
- b) Risk Communications Packet (to be submitted to CPS once external peer review has been completed.)

### 2) List of Presentations:

- a) EVENT: Center for Produce Safety, 2012 Produce Research Symposium  
TITLE: Assessment of *Escherichia Coli* as an Indicator of Microbial Quality of Irrigation Waters used for Produce  
DATE: June 27, 2012  
LOCATION: University of California, Davis, CA
- b) EVENT: Third Annual Food Safety Conference and Poster Session  
TITLE: Assessment of *Escherichia Coli* as an Indicator of Microbial Quality of Irrigation Waters used for Produce  
DATE: October 12, 2012  
LOCATION: Omni Tucson National, 2727 West Club Drive, Tucson AZ
- c) EVENT: Institute of the Environment Presents: Grad Blitz  
TITLE: Assessment of *Escherichia Coli* as an Indicator of Microbial Quality of Irrigation Waters used for Produce  
DATE: November 8, 2012  
LOCATION: Tucson Marriott, University Park, Tucson AZ  
AWARDS: Audience Choice Best Poster
- d) EVENT: Graduate and Professional Student Council Presents: Student Showcase  
TITLE: Assessment of *Escherichia Coli* as an Indicator of Microbial Quality of Irrigation Waters used for Produce  
DATE: November 9, 2012  
LOCATION: University of Arizona, University Mall, Tucson AZ

Assessment of *E. coli* as an Indicator of Microbial  
Quality for Irrigation Water Use for Produce

Project Funded by  
Center for Produce and Safety

Statistical Methods and Analysis  
Quantitative Microbial Risk Assessment  
(Technical Report)

Prepared  
By

Akrum H. Tamimi, Ph.D.  
[akrumt@email.arizona.edu](mailto:akrumt@email.arizona.edu)

and

Charles P. Gerba, Ph.D.  
[gerba@ag.arizona.edu](mailto:gerba@ag.arizona.edu)

Department of Soil, Water and Environmental Science  
429 Shantz Building #38, 1177 East Fourth St.  
The University of Arizona  
Tucson, AZ 85721-0038

January 31, 2013

## Table of Contents

Table of Contents .....	2
I. Distribution of <i>E. coli</i> Measurements .....	3
II. Statistical Analysis.....	7
A. Testing Significant Differences between the 3 Lab Methods .....	7
B. Testing Significant Differences between the 3 Locations.....	11
C. Analysis of Variance using Bootstrapping Technique.....	15
D. Testing Significant Differences between Methods at Each Location .....	15
1. Imperial Valley <i>E. coli</i> Counts .....	15
2. Maricopa <i>E. coli</i> Counts .....	18
3. Yuma <i>E. coli</i> Counts.....	21
III. Quantitative Microbial Risk Assessment.....	24
A. Dose-response Model for <i>E. coli</i> Infection .....	24
1. Exponential Model .....	24
2. Beta-Poisson Dose-Response Model.....	25
B. Evaluating Probability of Risk of Infection for Field Measured <i>E. coli</i> Concentrations	27
C. Results .....	28
IV. Epidemiology Based Risk Assessment.....	31
V. References.....	33

### I. Distribution of *E. coli* Measurements

*E. coli* counts in irrigation waters of irrigation canals in the Yuma area are shown in Figure 1, while Figures 2 and 3 show the *E. coli* counts for the Imperial Valley and Maricopa City in the Gila River Valley respectively. Figure 4 shows the *E. coli* counts determined for all sites. The maximum value plotted in Figures 1 through 4 is the maximum count obtained for the three different enumeration methods for each sample. The 126 level is shown as a line passing through the value of 126 *E. coli* count per 100 ml.

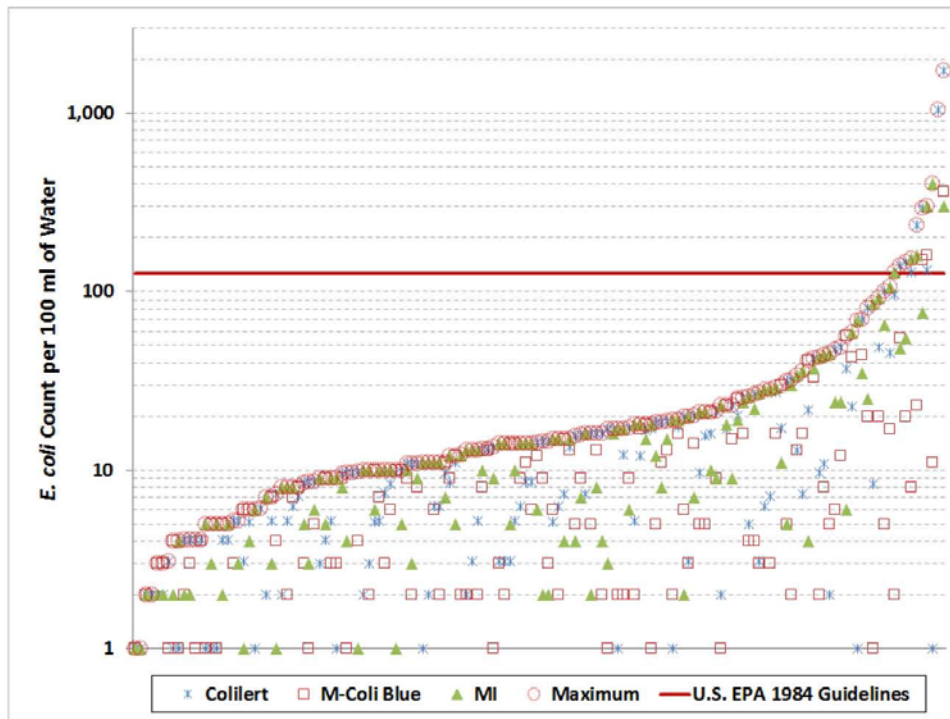


Figure 1: *E. coli* Counts for the Different lab Methods for Samples Obtained from Yuma

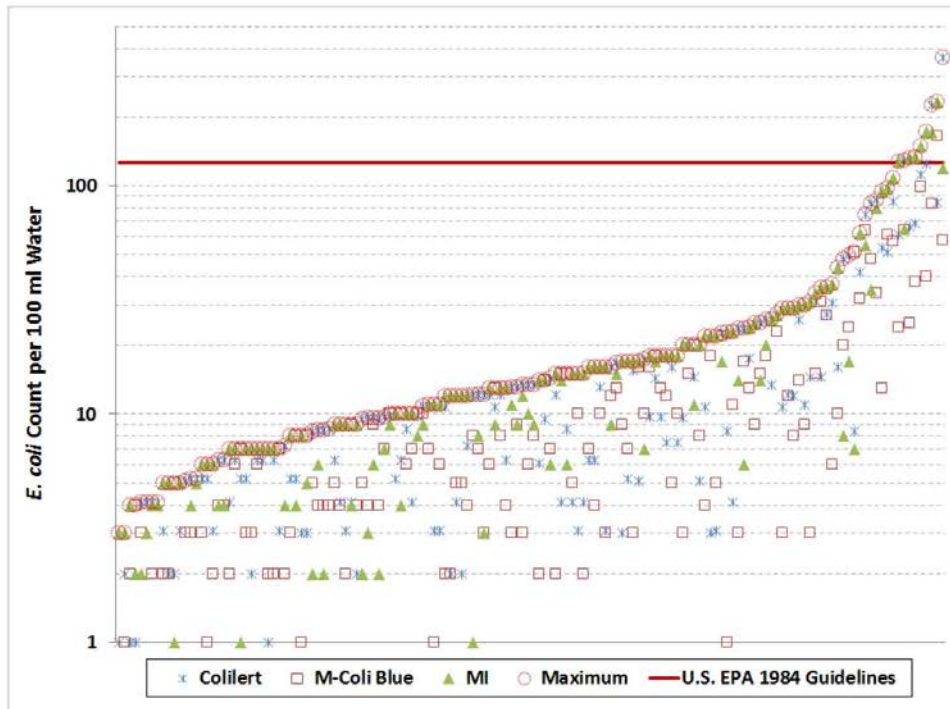


Figure 2: *E. coli* Counts for the Different lab Methods for Samples Obtained from Imperial Valley

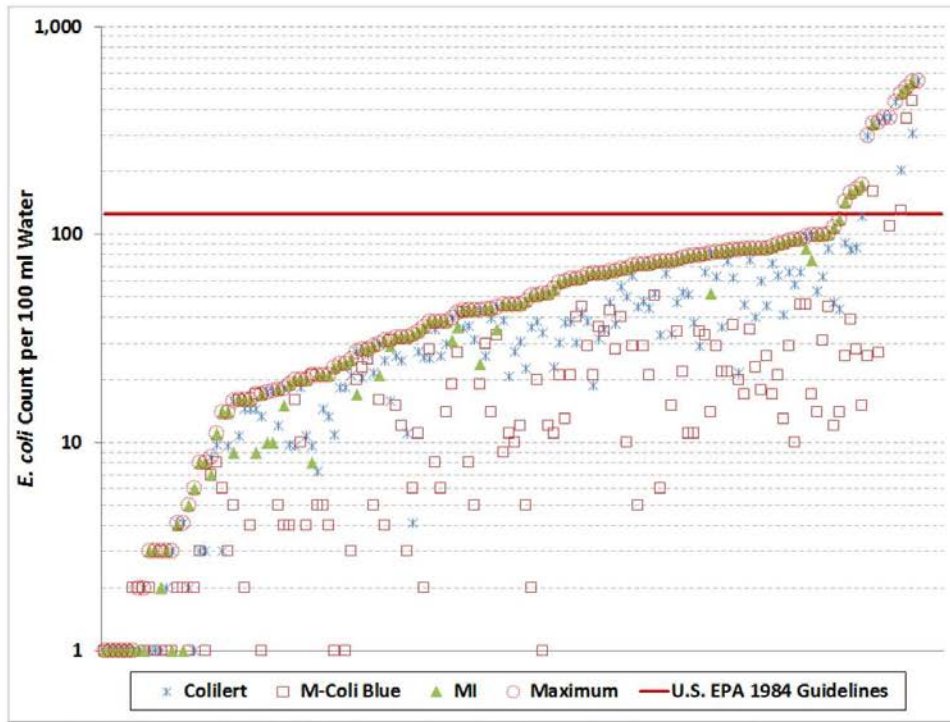


Figure 3: *E. coli* Counts for the Different lab Methods for Samples Obtained from Maricopa

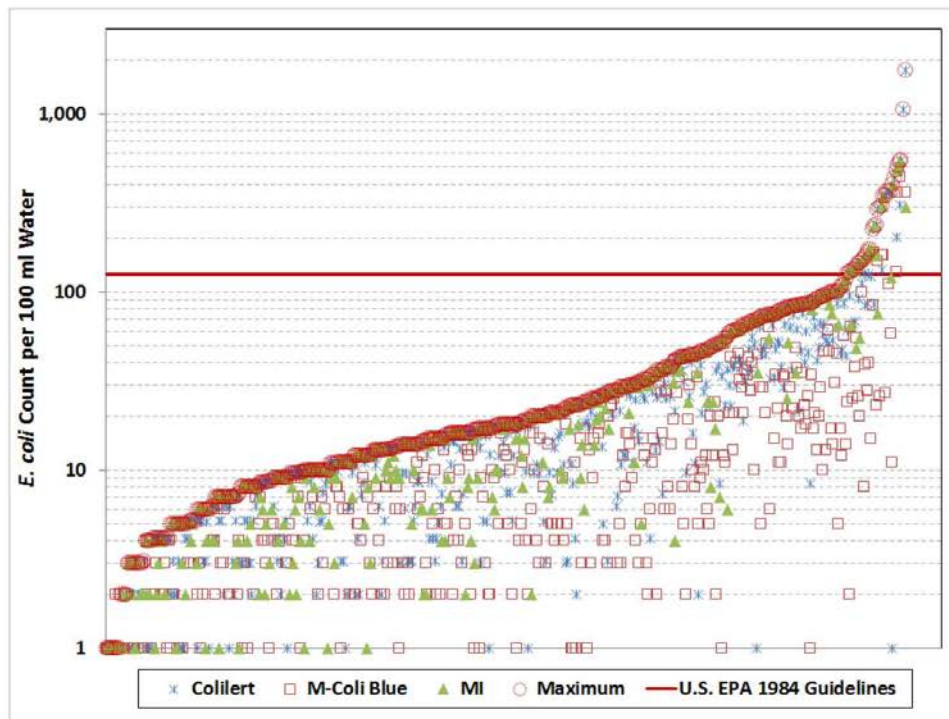


Figure 4: *E. coli* Counts for the Different lab Methods for Samples Collected from all Locations

## II. Statistical Analysis

All statistical and data analysis were conducted using the different packages in the programming Language R (The R Project for Statistical Computing, 2012).

### A. Testing Significant Differences between the 3 Lab Methods

To determine if there is statistical significant difference between the measurements obtained using the three different *E. coli* detection methods for all locations, analysis of variance (ANOVA) was conducted using the Language R (The R Project for Statistical Computing, 2012).

The actual values for *E. coli*/100 ml obtained from the different methods were used to conduct the ANOVA. Table 1 shows the different statistics obtained from the different methods used to test for *E. coli*.

Table 1: Statistics for the Different lab Methods

Statistics	Method		
	Colilert	M-Coli Blue	MI
Number of Measurements (N)	446	438	439
Arithmetic Mean	37.81	16.18	36.45
Standard Deviation	113.44	37.26	62.82

The condition of normality of the measured data was tested and the Q-Q plot was developed as shown in Figure 5. Figure 5 shows that the data did not come from a normally distributed population since most of the data lies outside the dotted lines on the graph representing the 5<sup>th</sup> and 95<sup>th</sup> percentiles.

It was determined to transform the *E. coli* counts using the  $\log_{10}$  and repeat the analysis on the transformed data.

Table 2 shows the statistics for the *E. coli* counts after applying the  $\log_{10}$  transformation to the data.

Table 2: Statistics for the Different Methods Based on  $\log_{10}$  Transformation of Measurements

Statistics	Method		
	Colilert E. Coli	M-Coli Blue	MI
Number of Measurements (N)	446	438	439
Geometric Mean	12.78	6.97	15.25
Standard Deviation of $\log_{10}$ Transformed Data	0.5932	0.5403	0.5931

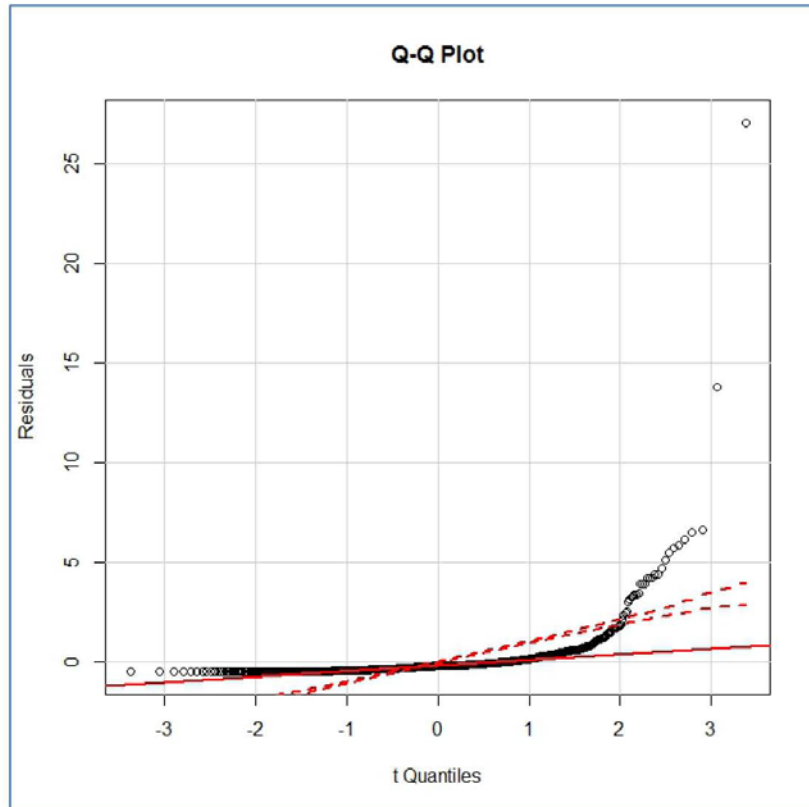


Figure 5: Q-Q Plot for *E. coli* Measurements using the Different Lab Methods

Figure 6 shows a plot of the statistics presented in Table 2 with 95% Confidence Interval (CI).

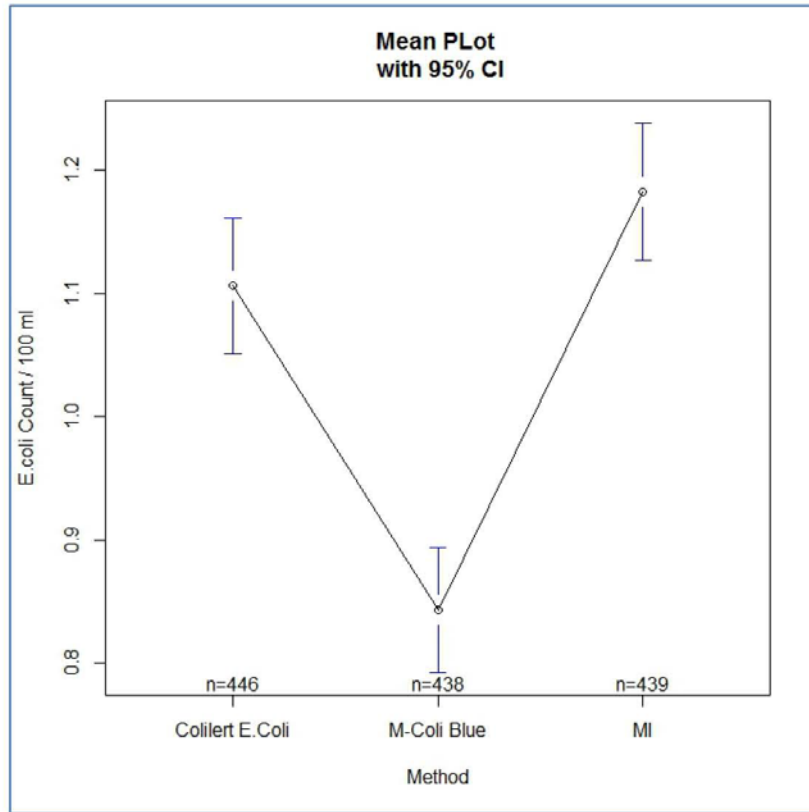


Figure 6: Geometric Mean Plot with 95% Confidence Interval (CI) for the *E. coli*  $\log_{10}$  Transformed Counts using the Different Lab Methods (n represents the number of measurements)

Figure 7 shows the Q-Q plot for the  $\log_{10}$  transformed *E. coli* counts. The plot shows that the transformed data has a distribution that is close to being normal and hence an assumption of normality is taken here.

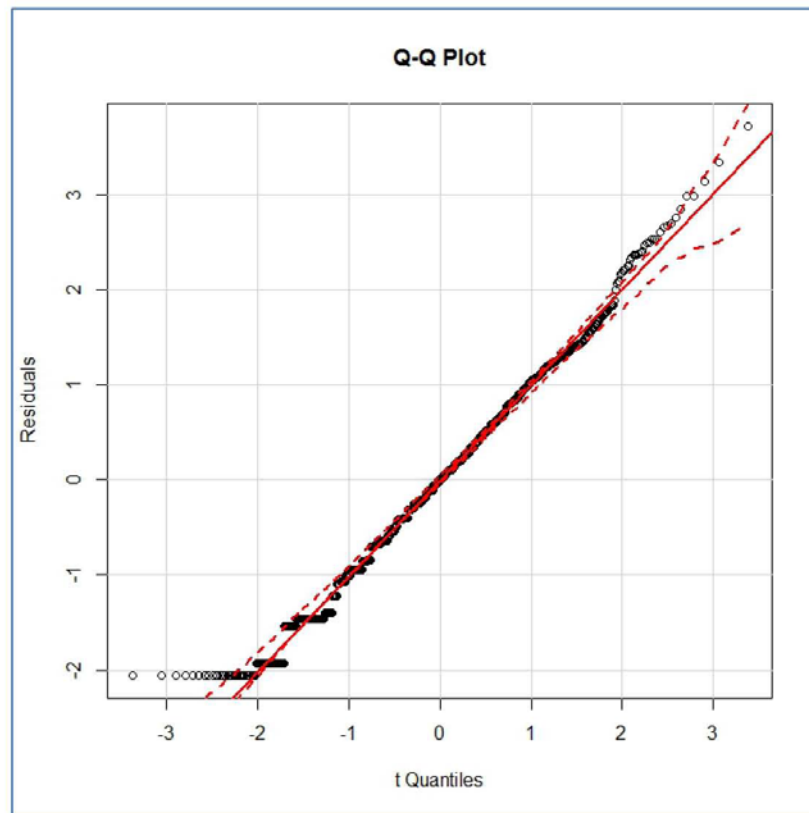


Figure 7: Q-Q Plot for *E. coli*  $\log_{10}$  Transformed Counts using the Different Lab Methods

To determine if the three lab methods have the same variance, one can conduct the Bartlett test of homogeneity of variances. The analysis produced a  $p$ -value of 0.08275 which is greater than 0.05 indicating that there are no significant differences between the variances of the 3 methods.

Finally, analysis of variance methodologies can be sensitive to the presence of outliers. One can test for outliers. The test indicates that there are no Studentized residuals with Bonferonni  $p$ -value of 0.2715 meaning that there are no outliers measured values among the three methods.

Since the normality test indicates that the  $\log_{10}$  transformed measured lab results for the 3 methods have normal distribution, equal variances and have no outlier values, then analysis of variance can be completed as shown below.

To determine if there is a significant difference in the *E. coli* counts per 100 ml of irrigation waters for the different testing methods, ANOVA was conducted and a  $p$ -value of  $2 \times 10^{-16}$  was determined indicating that there are significant differences between the different methods.

To compare each method to the other 2 methods, multiple comparison ANOVA tests were conducted and Table 3 shows the  $p$ -values.

**Table 3: Results of the ANOVA Multiple Comparison Tests**

<b>Method 1</b>	<b>Method 2</b>	<b><i>p</i>-value</b>
M-Coli Blue	Colilert	<b><i>&lt; 0.0001</i></b>
MI	Colilert	0.1160
MI	M-Coli Blue	<b><i>&lt; 0.0001</i></b>

*Note: Italicized and bold p-values indicate statistical significant difference*

The  $p$ -values in Table 3 show that there are significant differences between *E. coli* counts measured using M-Coli Blue and those measured using Colilert; and between those measured using Colilert and MI methods. However, with  $p$ -value of 0.1160 there are no significant differences between *E. coli* counts measured using MI and Colilert

### **B. Testing Significant Differences between the 3 Locations**

To determine if there is statistical significant differences between the measurements obtained at the 3 different locations being Yuma, Imperial Valley and Maricopa, analysis of variance (ANOVA) was conducted using the Language R.

Measured *E. coli* counts/100 ml of irrigation water obtained from the 3 different locations were used to conduct the ANOVA. Table 4 shows the different statistics obtained from the 3 locations.

**Table 4: Statistics for the Locations**

<b>Statistics</b>	<b>Location</b>		
	<b>Imperial Valley</b>	<b>Maricopa</b>	<b>Yuma</b>
<b>Number of Measurements (N)</b>	450	429	444
<b>Arithmetic Mean</b>	18.71	47.57	25.05
<b>Standard Deviation</b>	33.84	79.33	103.91

The condition of normality of the measured data at the different locations was tested and the Q-Q plot was developed and is presented in Figure 8. It can be seen that the normality condition is

not met indicating that the original measurements don't come from normally distributed population.

The measured *E. coli* count per 100 ml was transformed taking the  $\log_{10}$  of the measured counts per 100 ml and Table 5 shows the statistics for the 3 locations.

Figure 9 shows a plot of the statistics shown in Table 5 with 95% confidence intervals.

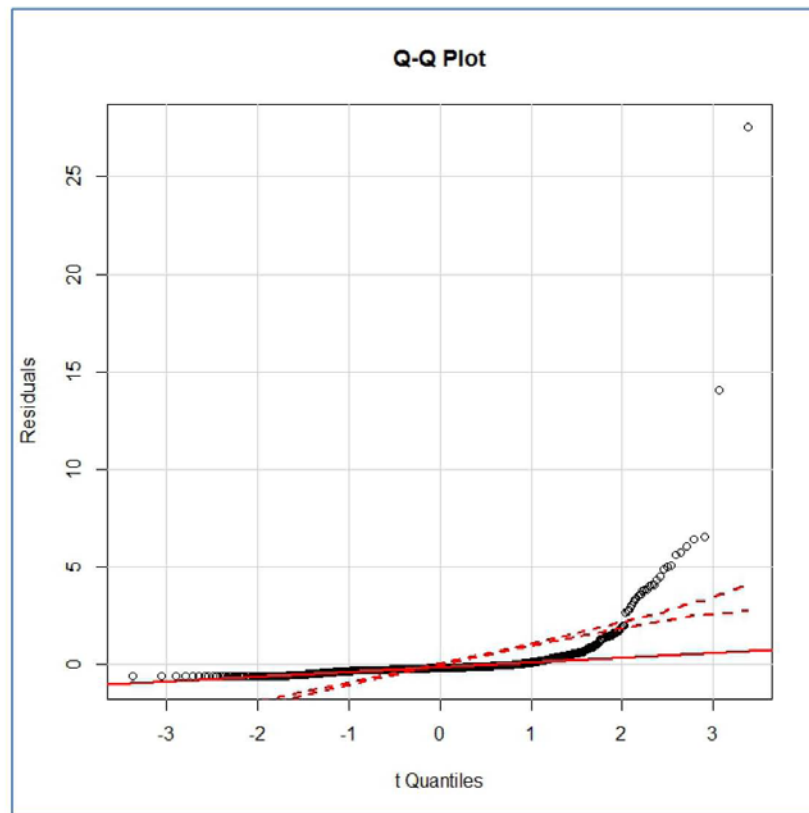
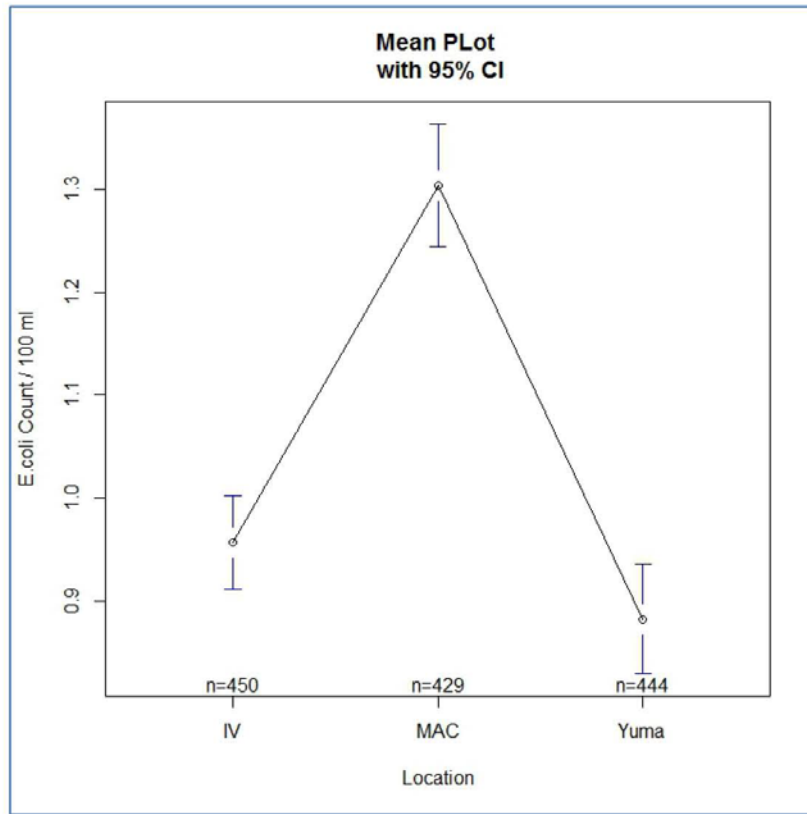


Figure 8: Q-Q Plot for *E. coli* Measurements at the Different Locations

**Table 5: Statistics for *E. coli* log<sub>10</sub> Transformed Counts at the different Locations**

Statistics	Location		
	Imperial Valley	Maricopa	Yuma
Number of Measurements (N)	450	429	444
Geometric Mean	9.06	20.15	7.63
Standard Deviation of log <sub>10</sub> Transformed Data	0.4901	0.6286	0.5726



**Figure 9: Geometric Mean Plot with 95% Confidence Interval (CI) for the *E. coli* log<sub>10</sub> Transformed Counts Grouping Different Locations- IV: Imperial Valley, MAC: Maricopa (n represents the number of measurements)**

Figure 10 shows the Q-Q plot for the  $\log_{10}$  transformed *E. coli* counts for the different locations. The plot shows that the  $\log_{10}$  transformed data has a distribution that can be approximated as being normally distributed and hence the assumption of normality is reasonable.

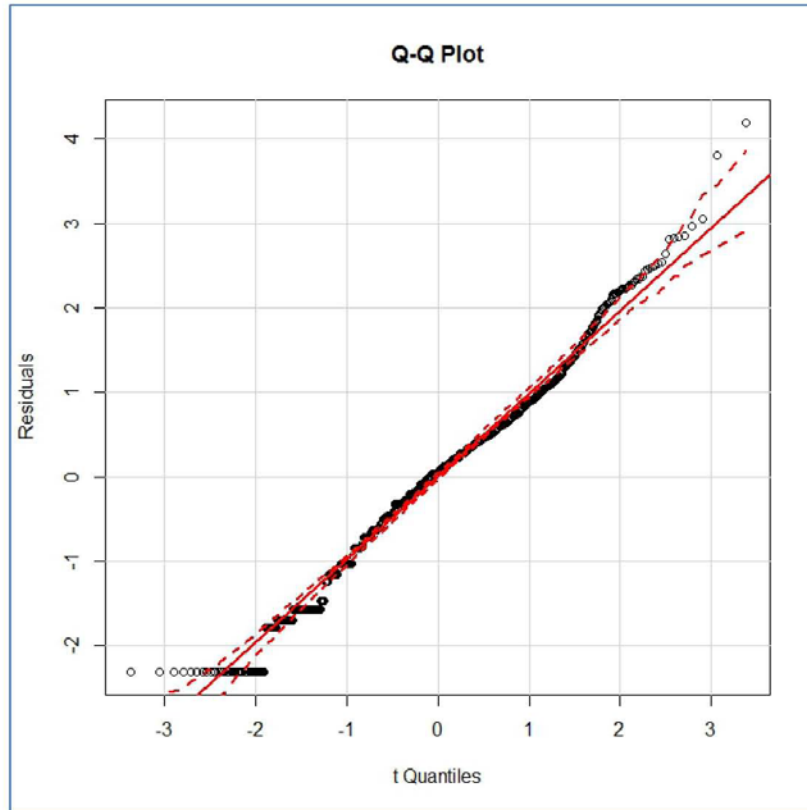


Figure 10: Q-Q Plot for *E. coli*  $\log_{10}$  Transformed Counts at the Different Locations

To determine if the *E. coli*  $\log_{10}$  transformed counts at the three locations have the same variance, one can run the Bartlett test of homogeneity of variances. The analysis produced a  $p$ -value less than 0.0001 indicating that there are significant differences in the variances between the 3 locations.

Finally, analysis of variance methodologies can be sensitive to the presence of outliers. One can test for outliers. The test indicates that there are Studentized residuals equals to 4.2 with Bonferonni  $p$ -value of 0.0381 meaning that there is one outlier measured value that corresponds to a Yuma reading of 1732.9 counts per 100 ml of irrigation water using the Colilert method.

Since there are significant differences amongst variances and there is an outlier value, one cannot perform the traditional ANOVA because 2 of the required 3 conditions are not met as described above.

To determine if there is a significant difference in the *E. coli* log<sub>10</sub> transformed counts for the different locations ANOVA was conducted and a *p*-value < 0.0001 was determined indicating that there are significant differences in the *E. coli* log<sub>10</sub> transformed counts between the different locations. Analysis of Variance using Bootstrapping Technique was used as shown below.

### C. Analysis of Variance using Bootstrapping Technique

Since the equal variances assumption is not true for the 3 locations and the data contains outliers, bootstrapping techniques are used here to determine the significant differences in the measured *E. coli* counts per 100 ml of irrigation water at the 3 locations.

Using 5000 bootstrap iterations to determine if there are significant differences between the 3 locations, analysis of variance resulted in a *p*-value < 0.0001 indicating that there are significant differences between the data measured at the 3 locations.

To compare each location to the other 2 locations, multiple comparison Bootstrapping ANOVA was conducted and table 6 shows the *p*-values for those tests indicating that there is significant differences between Maricopa and Imperial Valley and between Maricopa and Yuma with a *p*-value less than 0.00001 while Table 6 shows that there is no significant differences in the *E. coli* counts between Yuma and Imperial Valley locations.

**Table 6: Results of the ANOVA Multiple Comparison Tests for the Different Locations**

<b>Location 1</b>	<b>Location 2</b>	<b><i>p</i>-value</b>
Maricopa	Imperial Valley	<b><i>&lt; 0.00001</i></b>
Yuma	Imperial Valley	0.1200
Yuma	Maricopa	<b><i>&lt; 0.00001</i></b>

*Note: Italicized and bold p-values indicate statistical significant difference*

### D. Testing Significant Differences between Methods at Each Location

Since the log<sub>10</sub> transformed measured *E. coli* counts seem to fulfill the ANOVA requirements and assumptions, the statistical analysis in this section will only use the log<sub>10</sub> transformed measurements. However, the check for normality, homogeneity of the variances and the outliers test will be conducted for assurances. If it is proven that the measurements don't fulfill the ANOVA requirements, then bootstrapping techniques will be used for the analysis of variance.

#### 1. Imperial Valley *E. coli* Counts

The statistics of *E. coli* counts measured at Imperial Valley for the 3 different methods are shown in Table 7.

**Table 7: Statistics of *E. coli* measured counts per Method at Imperial Valley**

Statistics	Method		
	Colilert	M-Coli Blue	MI
Number of Measurements (N)	150	150	150
Arithmetic Mean	20.22	12.79	23.13
Standard Deviation	40.46	20.29	36.68

Table 8 shows the statistics obtained for the  $\log_{10}$  transformed measurements for each of the method at Imperial Valley.

**Table 8: Statistics of  $\log_{10}$  Transformed *E. coli* measured counts per Method at Imperial Valley**

Statistics	Method		
	Colilert	M-Coli Blue	MI
Number of Measurements (N)	150	150	150
Geometric Mean	9.57	6.65	11.68
Standard Deviation of $\log_{10}$ Transformed Data	0.4755	0.4790	0.4871

Figure 11 shows the mean plot for *E. coli* counts obtained at Imperial Valley for the three methods obtained from Table 8 with 95% confidence intervals for each method.

Figure 12 shows the Q-Q plot obtained from the R Language for determining the normality of the distribution of the data collected at Imperial Valley for the 3 lab methods. The figure shows that the  $\log_{10}$  transformed data follows approximate normal distribution since most of the measurements fall between the 2.5% and 97.5% percentiles.

The Bartlett test of homogeneity of variances provided p-value of 0.9555 indicating that there are no significant differences between the variances of the 3 lab methods conducted at Imperial Valley.

The outlier Test conducted on the Imperial Valley measurements for the 3 different methods indicated that there no Studentized residuals with Bonferonni p-value of 0.40908 indicating that there are no outliers.

Therefore, the ANOVA test was conducted on the  $\log_{10}$  transformed measured counts at Imperial Valley for the 3 methods. The p-value of  $5.71 \times 10^{-5}$  was obtained indicating that there are significant differences between the 3 methods.

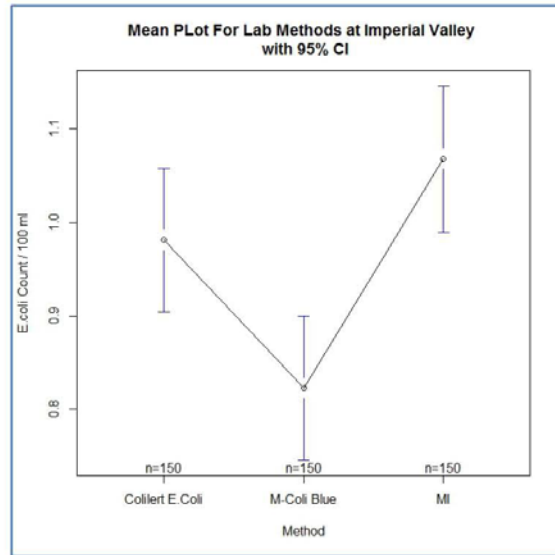


Figure 11: Mean Plot for the 3 Lab Methods at Imperial Valley

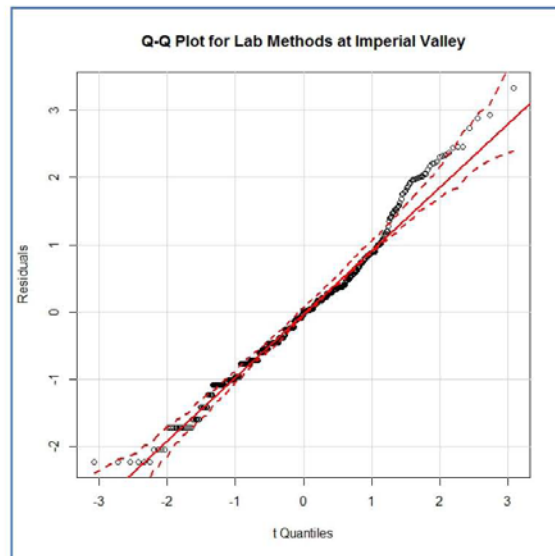


Figure 12: the Q-Q Plot Used for Normality Test

To determine which of the methods have significant statistical differences, multiple comparison ANOVA test was conducted resulting in the *p*-values shown in Table 9.

**Table 9: Multiple Comparison ANOVA Results for the 3 Methods Used at Imperial Valley**

<b>Method 1</b>	<b>Method 2</b>	<b><i>p</i>-value</b>
M-Coli Blue	Colilert	<b><i>0.0128</i></b>
MI	Colilert	0.2639
MI	M-Coli Blue	<b><i>0.00004</i></b>

*Note: Italicized and bold p-values indicate statistical significant difference*

The *p*-values shown in Table 9 indicate that there are significant differences between the *E. coli* measurements obtained from M-Coli Blue and Colilert *E. coli* methods; and between *E. coli* measurements obtained from lab methods MI and M-Coli Blue. However, the *p*-value for MI and Colilert *E. coli* methods showed that there are no significant differences between the measured values obtained by the 2 methods.

## 2. Maricopa *E. coli* Counts

The statistics of *E. coli* counts measured at Maricopa for the 3 different lab methods are shown in Table 10.

**Table 10: Statistics of *E. coli* Measured Counts per Method at Maricopa**

<b>Statistics</b>	<b>Method</b>		
	<b>Colilert</b>	<b>M-Coli Blue</b>	<b>MI</b>
<b>Number of Measurements (N)</b>	146	141	142
<b>Arithmetic Mean</b>	55.83	23.55	62.92
<b>Standard Deviation</b>	92.15	50.40	83.46

Table 11 shows the statistics obtained for the Log<sub>10</sub> transformed measurements for each of the method at Maricopa.

**Table 11: Statistics of Log<sub>10</sub> Transformed *E. coli* measured counts per Method at Maricopa**

<b>Statistics</b>	<b>Method</b>		
	<b>Colilert</b>	<b>M-Coli Blue</b>	<b>MI</b>
<b>Number of Measurements (N)</b>	146	141	142
<b>Geometric Mean</b>	24.62	10.29	31.98
<b>Standard Deviation of log<sub>10</sub> Transformed Data</b>	0.6088	0.5715	0.6002

Figure 13 shows the mean plot for *E. coli* counts obtained at Maricopa for the three methods obtained from Table 11 with 95% confidence intervals for each method.

Figure 14 shows the Q-Q plot for determining the normality of the distribution of the data collected at Maricopa for the 3 lab methods. The figure shows that the  $\log_{10}$  transformed data does not follow normal distribution since a lot of the measurements fall outside the 2.5% and 97.5% quantiles.

The Bartlett test of homogeneity of variances provided p-value of 0.7365 indicating that there are no significant differences between the variances of the 3 lab methods conducted at Maricopa.

The outlierTest conducted on Maricopa measurements for the 3 different methods indicated that there are no Studentized residuals with Bonferonni p-value of  $< 0.05$  indicating that there are no outliers.

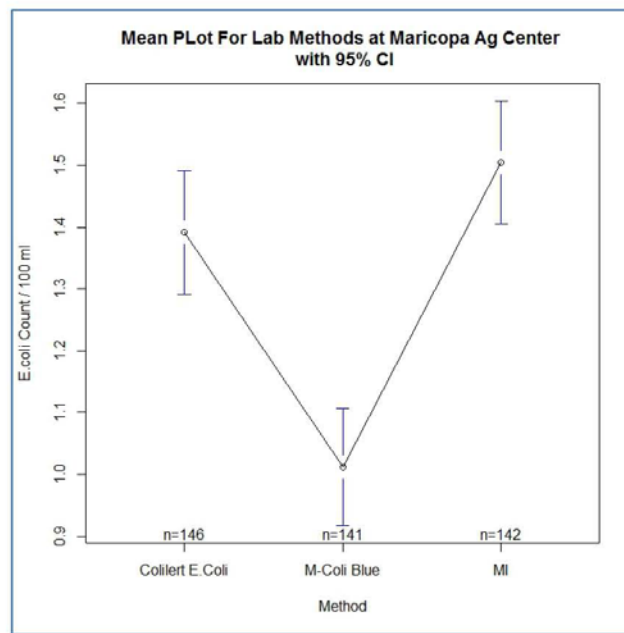


Figure 13: Mean Plot for the 3 Lab Methods at Maricopa

Since the normality assumption is not true for measurements obtained at Maricopa for the 3 lab methods, bootstrapping techniques are used to determine the significant differences in the measured *E. coli* counts per 100 ml of irrigation water utilizing the 3 lab methods at Maricopa.

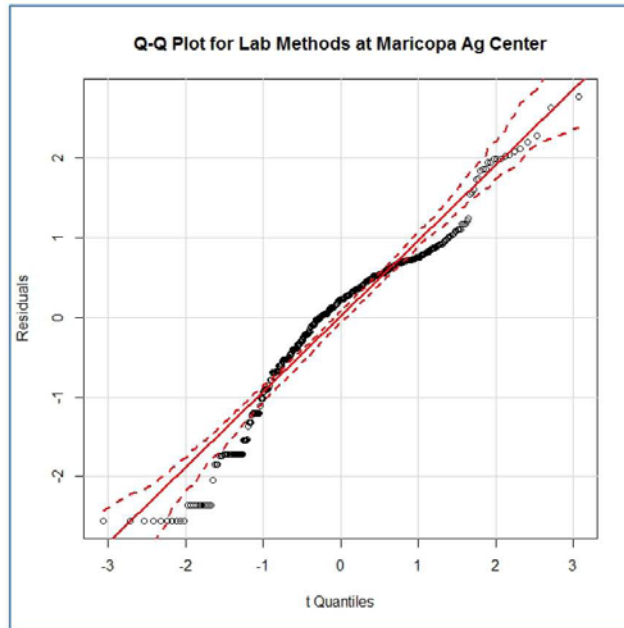


Figure 14: the Q-Q Plot Used for Normality Test for Measurements Obtained at Maricopa

Using 5000 bootstrap iterations to determine if there are significant differences between the 3 locations, analysis of variance resulted in a  $p$ -value  $2.2 \times 10^{-16}$  indicating that there are significant differences between the data measured using the different methods at Maricopa.

To compare each location to the other 2 locations, multiple comparison Bootstrapping ANOVA was conducted using 5000 iterations and table 12 shows the  $p$ -values for those tests indicating that there is significant differences between M-Coli Blue and Colilert lab testing methods used at Maricopa and MI and M-Coli Blue lab testing methods used at Maricopa. The  $p$ -value of 0.2371 shown in Table 12 suggests that there is no significant difference in the *E. coli* counts between MI and Colilert testing methods used at Maricopa.

Table 12: Multiple Comparison ANOVA Results for the 3 Methods Used at Maricopa

Method 1	Method 2	$p$ -value
M-Coli Blue	Colilert	< <b><i>0.00001</i></b>
MI	Colilert	0.2371
MI	M-Coli Blue	< <b><i>0.00001</i></b>

Note: *Italicized and bold p-values indicate statistical significant difference*

### 3. Yuma *E. coli* Counts

The statistics of *E. coli* counts measured at Yuma for the 3 different lab methods are shown in Table 13.

**Table 13: Statistics of *E. coli* Measured Counts per Method at Yuma**

Statistics	Method		
	Colilert	M-Coli Blue	MI
Number of Measurements (N)	150	147	147
Arithmetic Mean	37.86	12.56	24.48
Standard Deviation	167.05	34.97	52.27

Table 14 shows the statistics obtained for the Log<sub>10</sub> transformed measurements for each of the lab methods for Yuma measured data.

**Table 14: Statistics of Log<sub>10</sub> Transformed *E. coli* counts per Method at Yuma**

Statistics	Method		
	Colilert	M-Coli Blue	MI
Number of Measurements (N)	150	147	147
Geometric Mean	9.01	5.02	9.79
Standard Deviation of log <sub>10</sub> Transformed Data	0.5878	0.5271	0.5602

Figure 15 shows the mean plot for *E. coli* counts obtained at Yuma for the three methods with 95% confidence intervals for each method.

Figure 16 shows the Q-Q plot for determining the normality of the distribution of the data collected at Yuma for the 3 lab methods. The figure shows that the log<sub>10</sub> transformed data does not follow normal distribution since a lot of the measurements fall outside the 2.5% and 97.5% quantiles.

The Bartlett test of homogeneity of variances provided p-value of 0.4182 indicating that there are no significant differences between the variances of the 3 lab methods used for measurements in Yuma.

The outlierTest conducted on Yuma measurements for the 3 different methods indicated that there are no Studentized residuals with Bonferonni p-value of 0.0160 indicating that there are no outliers.

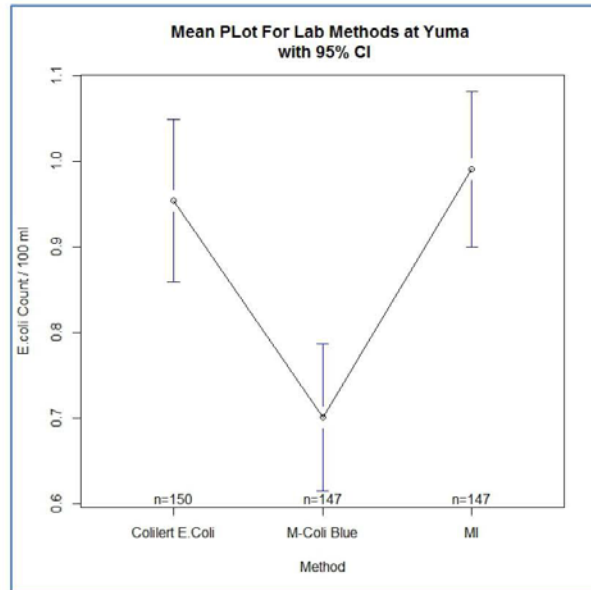


Figure 15: Mean Plot for the 3 Lab Methods Measured in Yuma

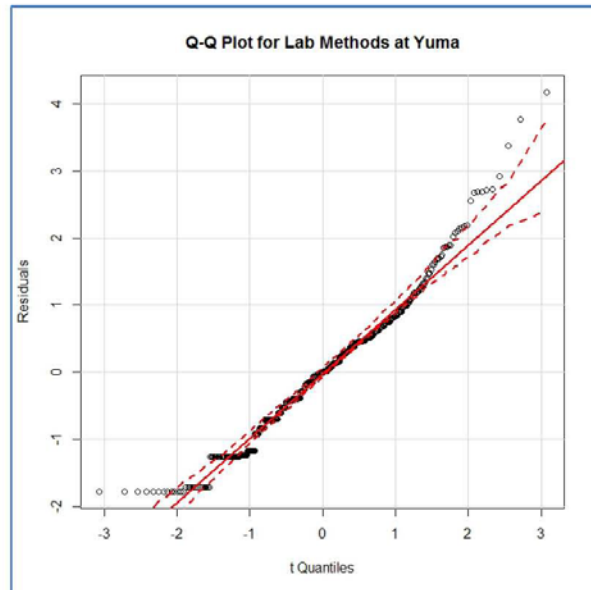


Figure 14: the Q-Q Plot Used for Normality Test for Measurements Obtained at Yuma

Since the normality assumption is not true for measurements obtained at Yuma for the 3 testing methods, bootstrapping techniques are used to determine the significant differences in the measured *E. coli* counts per 100 ml of irrigation water utilizing the 3 methods at Yuma.

Using 5000 bootstrap iterations to determine if there are significant differences between the 3 locations, analysis of variance resulted in a  $p$ -value  $2.2 \times 10^{-16}$  indicating that there are significant differences between the data measured using the 3 different methods at Yuma.

To compare each location to the other 2 locations, multiple comparison Bootstrapping ANOVA was conducted using 5000 iterations and table 15 shows the  $p$ -values for those tests indicating that there are significant differences between M-Coli Blue and Colilert lab testing methods used at Yuma and MI and M-Coli Blue lab testing methods used at Yuma. The  $p$ -value of 0.84159 shown in Table 12 suggests that there is no significant difference in the *E. coli* counts between MI and Colilert lab testing methods used in Yuma.

**Table 15: Multiple Comparison ANOVA Results for the 3 Methods Used at Yuma**

<b>Method 1</b>	<b>Method 2</b>	<b><i>p</i>-value</b>
M-Coli Blue	Colilert	<b><i>0.00032</i></b>
MI	Colilert	0.84159
MI	M-Coli Blue	<b><i>0.00003</i></b>

*Note: Italicized and bold p-values indicate statistical significant difference*

### III. Quantitative Microbial Risk Assessment

*Escherichia coli* (*E. coli*) O157:H7 is a food and waterborne pathogen that causes illness characterized by abdominal cramping, watery diarrhea followed by bloody diarrhea with little or no fever. In the following exercise all the *E. coli* detected are assumed to be pathogenic *E. coli* O157:H7. It is recognized that the detection methods used in this study do not detect *E. coli* O157:H7, but since *E. coli* O:157:H7 numbers are always lower in the water (Nwachuku and Gerba, 2008) then this would be a conservative assumption. A dose response was developed to estimate the risk of ingestion of *E. coli* from the irrigation water. This information was then used to model the risk from *E. coli* O157:H7 contamination of lettuce by three different irrigation methods i.e. sprinklers, furrow and subsurface drip.

#### A. Dose-response Model for *E. coli* Infection

To study the pathogenesis of diarrheal disease due to verotoxin (VT) producing *E. coli* O157:H7, Pai et al. (1986) infected 3-day-old rabbits with live *E. coli* O157:H7 and observed the effect of inoculum dose on diarrhea in the rabbits. Table 16 shows the dose response with probability of rabbits that suffered diarrhea.

Table 16: Dose-Response of Diarrhea in 3-day-old Rabbits

<i>E. coli</i> O157:H7 Dose (CFU*)	Number of Rabbits Inoculated	Number of Rabbits with Diarrhea	Observed Probability of Diarrhea
100,000	3	1	0.33
1,000,000	5	2	0.40
10,000,000	5	5	1.00
100,000,000	13	12	0.92
1,000,000,000	5	5	1.00
3,000,000,000	2	2	1.00
10,000,000,000	6	6	1.00

\*colony forming units

It was found that probability of infection for most microbial infections follows either an exponential model or the Beta-Poisson dose-response model (Haas, et al. 1999). To determine which model best fit the experimental data for *E. coli* O157:H7, the Solver routine in Microsoft Excel program was used to determine the best fit.

#### 1. Exponential Model

The exponential model can be expressed as shown in equation 1.

$$R = 1 - EXP(-k \times D) \tag{1}$$

In which,

- R*: is the probability of risk of infection, Response, (decimal);
- k*: is the infectivity per viable *E. coli* concentration which is derived from dose response studies (per colony forming units);
- D*: is the Dose of viable *E. coli* that is ingested (colony forming units or Organism).

The Solver routine in Microsoft Excel program was used to optimize the exponential model by varying the *k* value in equation 1 with the objective being maximizing the chi-square goodness of fit between the observed and the expected probability of risk of infection. The chi-square goodness-of-fit test is used due to the fact that the experiment is multinomial.

The null hypothesis in the chi-square goodness of fit test states that there is no statistical significant difference between the observed and the expected probability values for the respective dose; while the research hypothesis assumes that there is a significant statistical difference in at least one observed and expected probability value for the respective dose. The test statistic used is the  $\chi^2$ . The rejection region of the null hypothesis is defined based on an  $\alpha = 0.05$  and a degree of freedom of  $n - 1 = 7 - 1 = 6$ , where *n* is the number of observed probabilities.

Performing the optimization resulted in a *k* value of  $1.00 \times 10^{-6}$  per CFU and *p-value* of 0.99476 which indicates that there is no significant statistical difference between the observed and expected probabilities calculated using the exponential model.

Figure 15 shows a plot of the observed vs. the expected probabilities based on the exponential model.

## 2. Beta-Poisson Dose-Response Model

The Beta-Poisson model can be estimated by equation 2, in which  $\alpha$  and  $\beta$  are constants and *P* is the probability.

$$P(\text{Response}) = 1 - \left(1 + \frac{D}{\beta}\right)^{-\alpha} \quad (2)$$

Again, the Solver routine in Microsoft Excel program was used to optimize the Beta-Poisson dose response model by changing the values of  $\alpha$  and  $\beta$  in equation 2 with the objective being maximizing the chi-square goodness of fit between the observed and the expected probability of risk of infection.

Again, using the chi-square goodness of fit test with a  $\chi^2$  test statistic and a rejection region of the null hypothesis defined based on an  $\alpha = 0.05$  and a degree of freedom of  $n - 1 = 7 - 1 = 6$ , an  $\alpha = 0.34269$  and a  $\beta = 5.6732 \times 10^4$  were determined for equation 2 with *p-value* of 0.99996. Figure 2 shows the plot of observed vs. expected probabilities using the Beta-Poisson model.

Since the *p-value* for the Beta-Poisson dose response model shows a stronger fit between the observed and expected values, the Beta-Poisson model was used to determine probability of infection for *E. coli*.

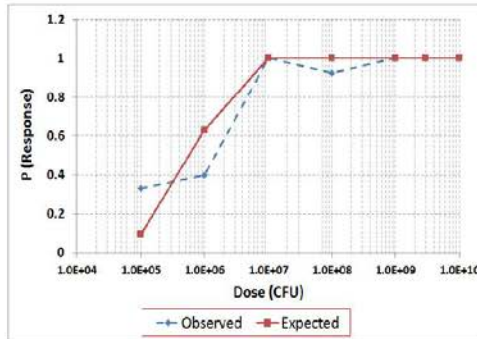


Figure 15: Optimization Output for Exponential Model

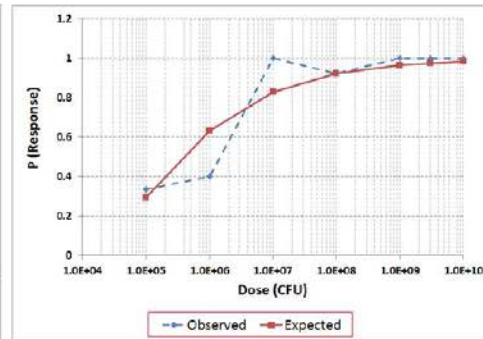


Figure 16: Optimization Output for Beta-Poisson Model

Figure 17 represents the Beta-Poisson dose-response curve that can be used to determine the probability of *E. coli* infection.

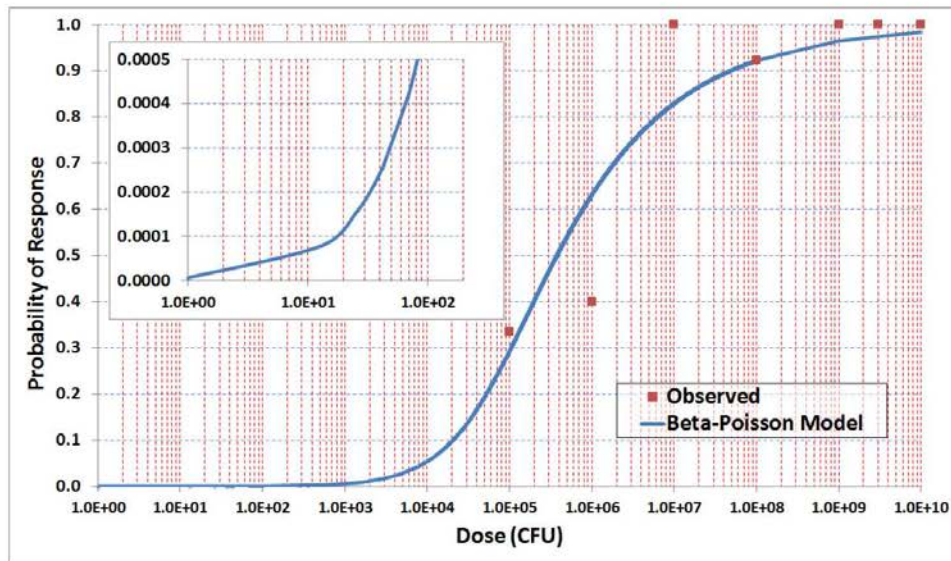


Figure 17: Dose Response for *E. coli* O157:H7

Figure 17 was developed based on the model shown in equation 3.

$$Probability (Infection) = 1 - \left( 1 + \frac{D}{5.6732 \times 10^4} \right)^{-(0.34269)} \quad (3)$$

### B. Evaluating Probability of Risk of Infection for Field Measured *E. coli* Concentrations

In this evaluation, the maximum *E. coli* concentrations measured in the field using the 3 different methods were utilized. The measured concentrations represent the irrigation water concentrations of *E. coli* per 100 ml of irrigation water. Concentration of 77 *E. coli* per 100 ml recommended by Johnson (2001), concentration of 126 *E. coli* per 100 ml recommended by U.S. EPA (1984) and concentration of 10,000 *E. coli* per 100 ml recommended by WHO (2006) were also utilized in calculating the risk of *E. coli* infections.

Figure 18 represents the annual probability of risk of infection for *E. coli* if a person directly ingested 100 ml of irrigation water on daily basis.

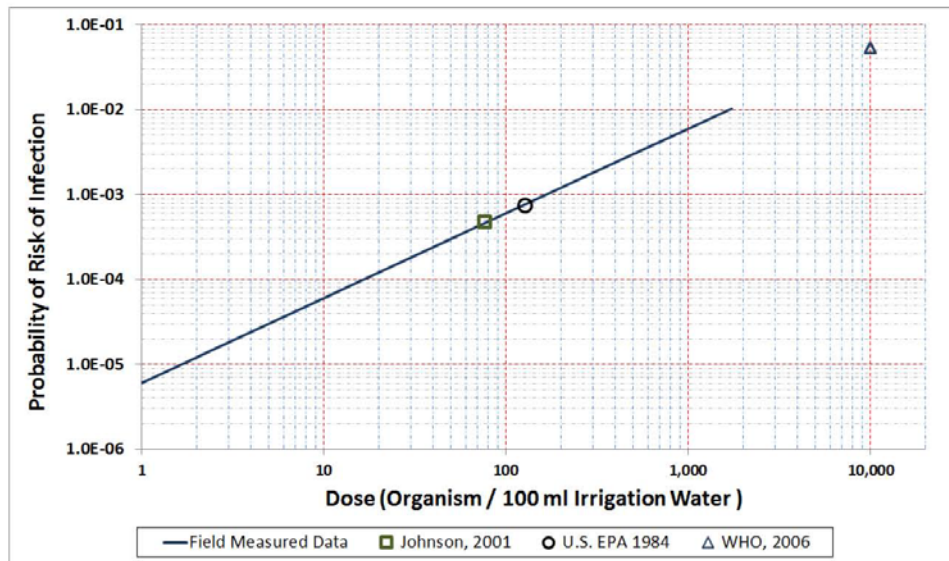


Figure 18: Probability of *E. coli* infection if a person ingested 100 ml of irrigation water on daily basis

Notice the legend in the graph for the different concentrations of *E. coli* per 100 ml of irrigation water used in the literature.

Stine et al. (2005) measured in the field the fraction of *E. coli* that is transferred to lettuce when irrigated with water inoculated with *E. coli*. The maximum ratio for furrow irrigation systems was found to be  $1.1 \times 10^{-4}$  and for subsurface irrigation systems to be  $8.8 \times 10^{-7}$ . Stine et al. (2011) found that the transfer fraction for sprinkler irrigation systems to be  $1.1 \times 10^{-2}$ . Stine et al. assumed in both the 2005 and 2011 studies that the adjusted per capita consumption of lettuce in the United States is 4,416.5 grams and that the consumption of lettuce occurs one day after irrigation.

Multiplying each fraction for the respective irrigation system by the maximum concentration of the *E. coli* measured in the field yields figures 19 through 21 that represent the annual risk of infection from eating an average of 12.1 grams of fresh lettuce per day when the lettuce is irrigated via the specific irrigation system. Notice the legend in the graph for the different concentrations of *E. coli* per 100 ml of irrigation water found in the literature that have been extensively used.

### C. Results

The results of risk assessment indicative that irrigation of lettuce with water containing 126 *E. coli* / 100 ml would result in a risk of less than 1:300,000 from a onetime consumption of the lettuce for all types of irrigation methods. The risks for subsurface drip irrigation were less than  $10^{-9}$ . Even considering multiple exposures (consumption of contaminated lettuce several times) these risks would be exceptionally low. It would appear based on these assumptions, the risk of transmission of *E. coli* 0157:H7 using irrigation water with 126 *E. coli* per 100 ml would be exceptionally low.

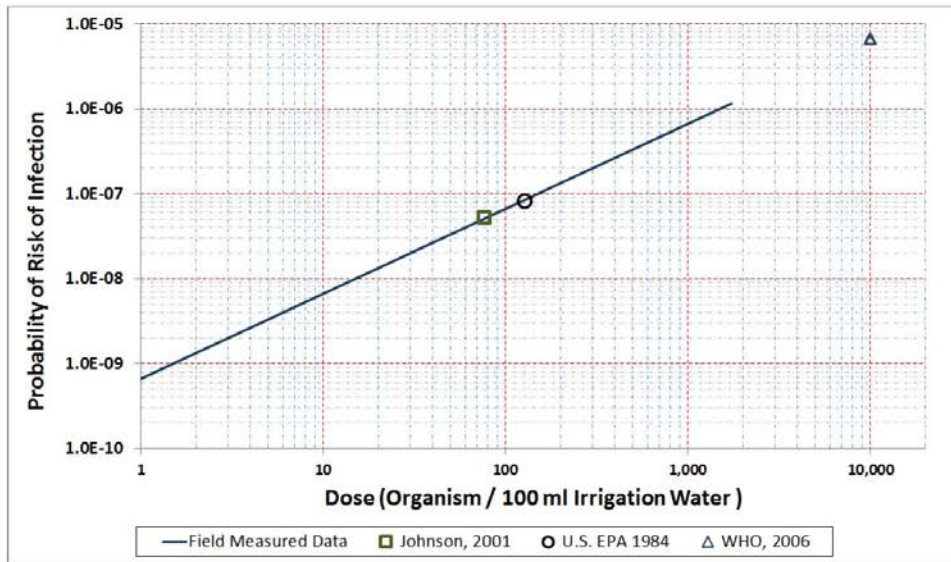


Figure 19: Annual Probability of *E. coli* Infection If a Person consumes 12.1 grams per day of Fresh Lettuce Irrigated Using a Furrow Irrigation System

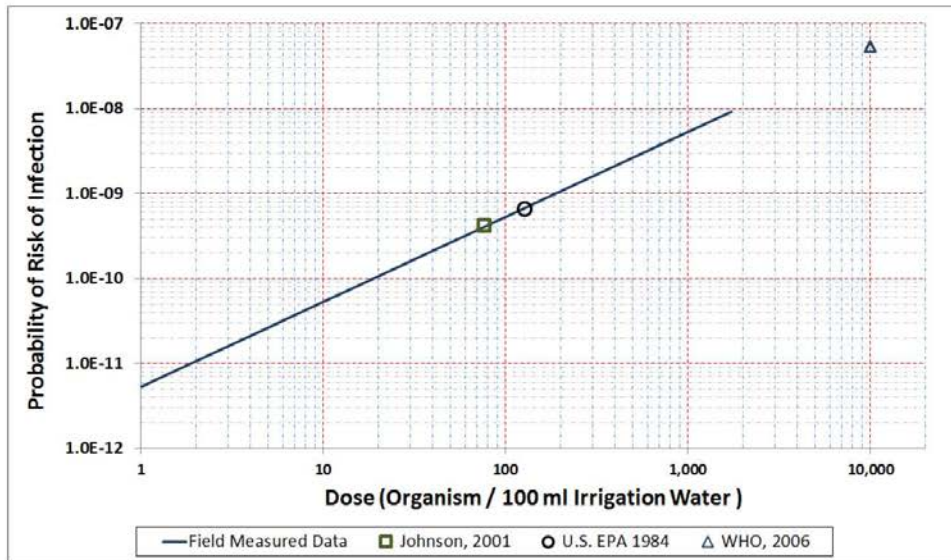


Figure 20: Annual Probability of *E. coli* Infection If a Person consumes 12.1 grams per day of Fresh Lettuce Irrigated Using a Subsurface Drip Irrigation System

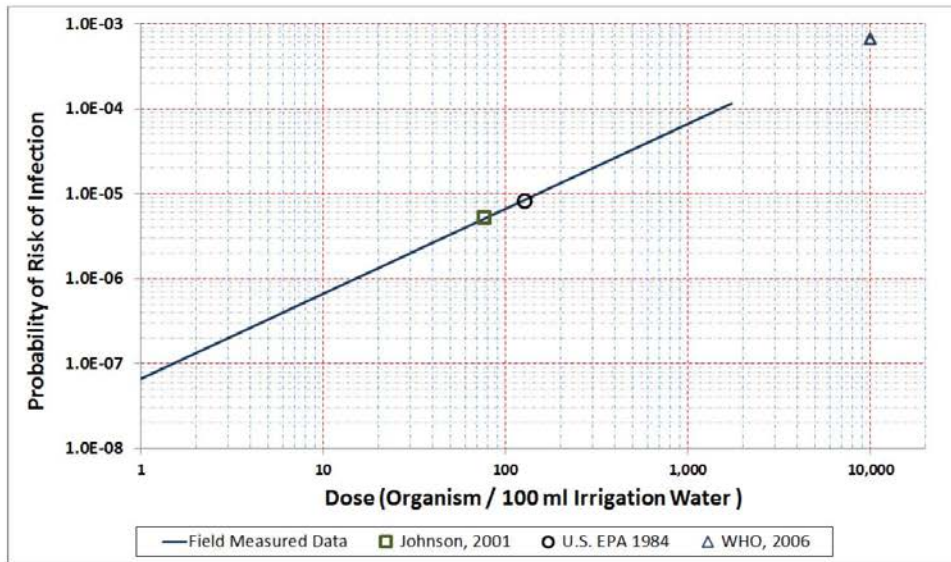


Figure 21: Annual Probability of *E. coli* Infection If a Person consumes 12.1 grams per day of Fresh Lettuce Irrigated Using a Sprinkler Irrigation System

#### IV. Epidemiology Based Risk Assessment

The U.S. EPA (1984) developed a health affects criteria for swimming in fresh recreational waters using *E. coli* as an indicator using epidemiological data on gastrointestinal illness (GI) among bathers. They suggested as a standard 126 *E. coli*/100ml as the recreational standard. They found that this corresponds to 8 illnesses per 1,000 bathers. The cause of the illness was not known, but believed to be of largely viral origin. All of the bathing waters studied where impacted by sewage discharges. Figure 22 adapted from U.S. EPA (1984) shows probability of gastrointestinal (GI) illness as a function of *E. coli* density per 100 ml of swimming water.

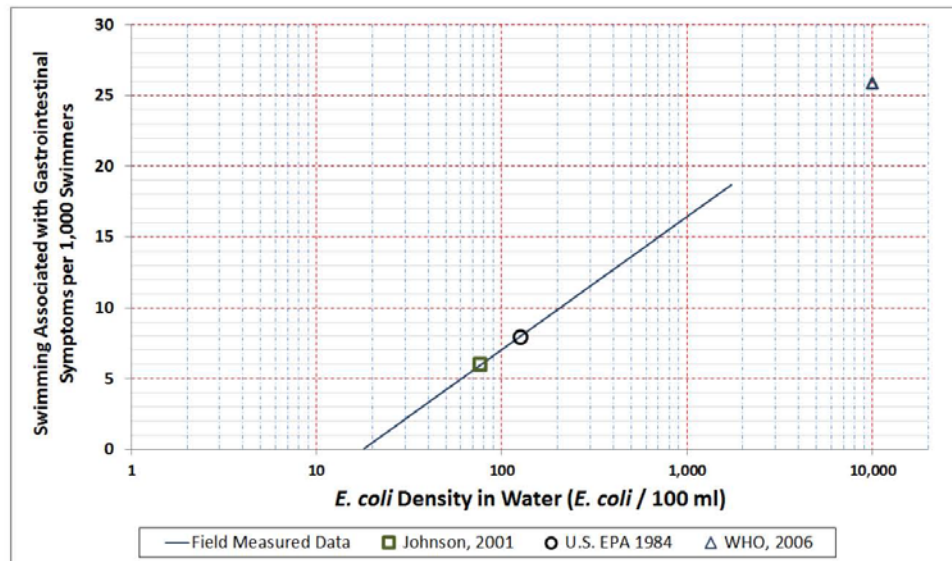


Figure 22: Criterion for Estimating Swimming-Associated Gastrointestinal Illness from *E. coli* Density per 100 ml in Fresh Recreational Water

Figure 22 shows that 8 in 1000 swimmers will have GI illness if they swim in fresh recreational water that has *E. coli* density of 126 per 100 ml. It is assumed that bathers swallow 10 ml or ingest 12.6 *E. coli* when they swim in water containing 12.6 *E. coli* and this will result in 8 GI illnesses per 1,000 persons.

Using a proportional linear approach with Figure 22, if irrigation water has *E. coli* density of 126 per 100 ml (or 12.6 *E. coli* per 10 ml), and based on Stine et al. (2005), 0.00011 of the 126 *E. coli* per 100 ml (or 12.6 *E. coli* per 10 ml) will be transferred to lettuce for furrow irrigation system and  $8.8 \times 10^{-7}$  of the 126 will be transferred to lettuce for subsurface drip irrigation

system. That corresponds to a risk of GI illness of 1.1 in 100,000 for furrows and 9 in 100,000,000 for subsurface irrigation system.

For sprinkler irrigation system and based on Stine et al. (2011), 0.011 of the 126 *E. coli* per 100 ml (or 12.6 *E. coli* per 10 ml) will be transferred to lettuce resulting in a risk of GI illness of 1.1 in 1,000.

Based on these assumptions the risk of illness from consumption of lettuce with irrigation water would be low. This assessment only applies to lettuce. Uncertainty in the estimate could be reduced by better data on transfer efficiency of spray irrigation since the data of Stine et al. (2011) was based on the use of pesticide spray from a container. Also, this assessment is only a onetime evident and multiple contamination or multiple days of consuming the same head of lettuce would increase the risk. Still irrigation water containing 126 *E. coli*/100 ml for lettuce would appear to present a minimal risk for furrow and subsurface drip. However, further research on contamination of lettuce by spray irrigation appears warrant to reduce uncertainty in the risk estimate.

## V. References

- Hass, C. N., J. B. Rose, and C. P. Gerba. 1999. Quantitative microbial risk assessment. John Wiley and Sons, Inc. New York.
- Johnson, T. R. (2001). Water Quality Criteria for Microbiological Indicators: Overview Report. <http://www.env.gov.bc.ca/wat/wq/BCguidelines/microbiology/microbiology.html#tab1> Accessed November 29, 2011. Last Modified 2001.
- Nwachuku, N. and C. P. Gerba. 2008. Occurrence and persistence of *Escherichia coli* 0157:H7. *Rev. Environ. Sci. Biotechnol.*, 7:267-273
- Pai, C.H., Kelly, J.K., Meyers, G.L., 1986. Experimental infection of infant rabbits with verotoxin-producing *E. coli*. *Infect. Immun.* 51 (1), 16–23.
- Stine, S. W., Song, I., Choi, C. Y., and Gerba, C. P. (2005). Application of microbial risk assessment to the development of standards for enteric pathogens irrigation water. *Journal of Food Protection*, 68, 913-918.
- Stine, S. W., Song, I., Choi, C. Y., and Gerba, C. P. (2011). Application of Pesticide Sprays to Fresh Produce: A risk assessment for Hepatitis A and Salmonella. *Food Environ Virol.* Springer Science and Business Media.
- The R Project for Statistical Computing (2012). <http://www.r-project.org/>
- U.S. EPA. 1984. Health Effects Criteria for Fresh Recreational Waters. EPA-600/1-84-004. Health Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency Research Triangle Park, NC 277711.
- WHO. (2006). World Health Organization and United Nations Environmental Programs. WHO Guidelines for the Safe Use of Wastewater, Excreta and Greywater. Geneva, Switzerland: WHO Press, 2006.

## Appendix B.

### 1) Budget Summary

Over the course of the project, grant funds were used to cover costs associated with the following:

- Salaries for research specialists and students. Responsibilities included collection of irrigation water samples from Maricopa and Yuma, AZ as well as Imperial Valley, CA. Additionally, these specialists and students were responsible for all method evaluation and molecular confirmatory work for all three methods and regions evaluated.
- Salary for research scientist. Responsibilities included assisting Dr. Gerba with the Quantitative Microbial Risk Assessment (QMRA) and data reporting where appropriate. Also, advising the project members on data interpretation.
- Travel. Costs were allocated in the budget for travel from Tucson and Maricopa, AZ to Yuma, AZ and Imperial Valley, CA. Costs included mileage, per diem for lodging and meals, as well as any incidentals incurred.
- Consumables. Finally, a substantial amount of the budget was dedicated to expendable laboratory supplies, water sample testing, consumables for molecular confirmatory work and sample shipping costs.
- It should be noted that while all work was completed within the budgetary limitations of the grant, in the future, the research team would propose to increase funds dedicated towards travel for sample collection, project meetings with PIs, and meetings with the industry. A significant amount of time was spent during the onset of the project working with the local industry to establish goals and sampling locations that would benefit not only the project, but also the local industry. Additionally, determining logistically the most feasible plan of action took multiple trips prior to the onset of sample collection.

## **Appendix C.**

- 1) Tables and Figures (See references in text to Appendix A.)
- 2) Raw Data Excel Plots for Sequencing Analysis

**Table 1.** Sequencing Results for MI Agar

PCR Number used for Submission	Isolate Number	Organism ID	Percent ID	FP/FN	Media Type
SAMP_118	A131	Enterobacter	43%	FP	MI
SAMP_129	A199	Aeromonas	100%	FP	MI
SAMP_133	A243	Enterobacter	60%	FP	MI
SAMP_134	A244	Escherichia/Shigella	100%	FP	MI
SAMP_136	A247	Enterobacter	73%	FP	MI
SAMP_137	A249	Enterobacter	53%	FP	MI
SAMP_138	A258	Klebsiella	97%	FN	MI
SAMP_15	A100	Enterobacter	85%	FP	MI
SAMP_17	A111	Pseudomonas	100%	FP	MI
SAMP_18	A116	Enterobacter	100%	FN	MI
SAMP_19	A157	Enterobacter	95%	FN	MI
SAMP_21	A181	Escherichia/Shigella	100%	FP	MI
SAMP_216	MAR-20	Escherichia/Shigella	100%	FP	MI
SAMP_22	A183	Enterobacter	97%	FP	MI
SAMP_221	MAR-01	Klebsiella	99%	FP	MI
SAMP_222	MAR-21	Providencia	100%	FP	MI
SAMP_23	A184	Enterobacter	100%	FP	MI
SAMP_24	A185	Escherichia/Shigella	100%	FP	MI
SAMP_240	MAR-18	Enterobacter	100%	FP	MI
SAMP_241	MAR-17	Escherichia/Shigella	100%	FP	MI
SAMP_242	MAR-36	Enterobacter	93%	FP	MI
SAMP_243	MAR-40	Enterobacter	84%	FP	MI
SAMP_245	MAR-41	Escherichia/Shigella	100%	FP	MI
SAMP_246	MAR-39	Enterobacter	98%	FP	MI
SAMP_249	MAR-25	Enterobacter	92%	FP	MI
SAMP_25	A197	Klebsiella	94%	FN	MI
SAMP_254	MAR-131	Escherichia/Shigella	100%	FP	MI
SAMP_255	MAR-137	Escherichia/Shigella	100%	FP	MI
SAMP_257	MAR-140	Acinetobacter	100%	FN	MI
SAMP_259	MAR-142	Escherichia/Shigella	100%	FP	MI

SAMP_260	MAR-144	Acinetobacter	100%	FP	MI
SAMP_261	MAR-174	Vibrio	100%	FP	MI
SAMP_262	MAR-176	Vibrio	100%	FP	MI
SAMP_27	A224	Escherichia/Shigella	100%	FP	MI
SAMP_278	MAR-07	Acinetobacter	100%	FN	MI
SAMP_279	MAR-09	Klebsiella	97%	FN	MI
SAMP_28	A226	Escherichia/Shigella	100%	FP	MI
SAMP_281	MAR-11	Acinetobacter	100%	FN	MI
SAMP_288	MAR-177	Enterobacter	86%	FP	MI
SAMP_289	MAR-181	Enterobacter	95%	FN	MI
SAMP_29	A227	Escherichia/Shigella	100%	FP	MI
SAMP_290	MAR-182	Enterobacter	84%	FN	MI
SAMP_296	MAR-19	Escherichia/Shigella	100%	FP	MI
SAMP_297	MAR-22	Enterobacter	98%	FP	MI
SAMP_298	MAR-29	Acinetobacter	100%	FN	MI
SAMP_299	MAR-32	Klebsiella	100%	FN	MI
SAMP_30	A229	Escherichia/Shigella	100%	FP	MI
SAMP_300	MAR-38	Enterobacter	89%	FP	MI
SAMP_303	A134	Enterobacter	49%	FN	MI
SAMP_304	MAR-08	Enterobacter	100%	FN	MI
SAMP_306	A108	Escherichia/Shigella	100%	FP	MI
SAMP_308	A117	Pseudomonas	100%	FN	MI
SAMP_314	MAR-15	Enterobacter	67%	FN	MI
SAMP_316	A139	Enterobacter	58%	FP	MI
SAMP_32	A233	Escherichia/Shigella	100%	FP	MI
SAMP_322	A142	Enterobacter	63%	FP	MI
SAMP_328	MAR-27	Enterobacter	97%	FP	MI
SAMP_331	A114	Klebsiella	100%	FN	MI
SAMP_34	A237	Pseudomonas	100%	FN	MI
SAMP_341	A173	Enterobacter	79%	FP	MI
SAMP_343	A177	Klebsiella	97%	FN	MI
SAMP_348	MAR-130	Enterobacter	99%	FN	MI
SAMP_349	A72	Enterobacter	95%	FP	MI
SAMP_35	A239	Enterobacter	91%	FN	MI

SAMP_350	A112	Enterobacter	98%	FP	MI
SAMP_41	A63B	Enterobacter	99%	FP	MI
SAMP_42	A64A	Enterobacter	93%	FP	MI
SAMP_44	A70	Enterobacter	100%	FP	MI
SAMP_45	A73	Pantoea	75%	FP	MI

**Table 2.** Sequencing Results for m-ColiBlue24® broth

PCR Number used for Submission	Isolate Number	Organism ID	Percent ID	FP/FN	Media Type
SAMP_148	B114	Enterobacter	91%	FN	MCOLIBLUE
SAMP_149	B139	Enterobacter	93%	FN	MCOLIBLUE
SAMP_150	B142	Enterobacter	70%	FP	MCOLIBLUE
SAMP_151	B150	Enterobacter	51%	FP	MCOLIBLUE
SAMP_154	B178	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_156	B187	Enterobacter	40%	FP	MCOLIBLUE
SAMP_160	B193	Enterobacter	94%	FN	MCOLIBLUE
SAMP_162	B231	Enterobacter	56%	FN	MCOLIBLUE
SAMP_164	B69	Enterobacter	92%	FP	MCOLIBLUE
SAMP_215	MAR-46	Klebsiella	99%	FN	MCOLIBLUE
SAMP_248	MAR-48	Enterobacter	97%	FP	MCOLIBLUE
SAMP_256	MAR-138	Enterobacter	91%	FP	MCOLIBLUE
SAMP_258	MAR-06	Enterobacter	100%	FP	MCOLIBLUE
SAMP_263	MAR-45B	Pseudomonas	100%	FN	MCOLIBLUE
SAMP_264	MAR-59	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_266	MAR-62	Enterobacter	90%	FP	MCOLIBLUE
SAMP_267	MAR-67	Klebsiella	99%	FN	MCOLIBLUE
SAMP_269	MAR-65	Enterobacter	98%	FN	MCOLIBLUE
SAMP_270	MAR-68	Klebsiella	98%	FN	MCOLIBLUE
SAMP_271	MAR-69	Enterobacter	93%	FN	MCOLIBLUE
SAMP_284	MAR-147	Klebsiella	100%	FP	MCOLIBLUE
SAMP_285	MAR-148	Enterobacter	100%	FP	MCOLIBLUE
SAMP_291	MAR-188	Acinetobacter	100%	FP	MCOLIBLUE
SAMP_292	MAR-187	Acinetobacter	100%	FP	MCOLIBLUE
SAMP_293	MAR-190	Escherichia/Shigella	98%	FP	MCOLIBLUE
SAMP_294	MAR-191	Enterobacter	98%	FP	MCOLIBLUE
SAMP_307	B71	Enterobacter	90%	FP	MCOLIBLUE

SAMP_309	B53	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_312	B72	Enterobacter	78%	FP	MCOLIBLUE
SAMP_315	B98	Klebsiella	84%	FN	MCOLIBLUE
SAMP_319	B64(2)	Klebsiella	100%	FN	MCOLIBLUE
SAMP_320	B52	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_323	MAR-143	Acinetobacter	100%	FP	MCOLIBLUE
SAMP_325	B120	Enterobacter	99%	FP	MCOLIBLUE
SAMP_332	B75	Enterobacter	94%	FP	MCOLIBLUE
SAMP_336	MAR-136	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_337	B128	Enterobacter	37%	FP	MCOLIBLUE
SAMP_340	B73	Enterobacter	100%	FP	MCOLIBLUE
SAMP_345	B69	Enterobacter	100%	FP	MCOLIBLUE
SAMP_46	B103	Enterobacter	81%	FP	MCOLIBLUE
SAMP_47	B104	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_48	B110	Klebsiella	100%	FP	MCOLIBLUE
SAMP_49	B111	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_50	B118	Klebsiella	100%	FN	MCOLIBLUE
SAMP_51	B140	Klebsiella	97%	FN	MCOLIBLUE
SAMP_52	B160	Enterobacter	93%	FP	MCOLIBLUE
SAMP_53	B164	Enterobacter	88%	FP	MCOLIBLUE
SAMP_54	B168	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_56	B173	Citrobacter	100%	FN	MCOLIBLUE
SAMP_57	B198	Enterobacter	91%	FP	MCOLIBLUE
SAMP_58	B199	Enterobacter	99%	FP	MCOLIBLUE
SAMP_59	B202	Enterobacter	58%	FP	MCOLIBLUE
SAMP_60	B203	Enterobacter	84%	FP	MCOLIBLUE
SAMP_61	B207	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_63	B212	Enterobacter	70%	FN	MCOLIBLUE
SAMP_64	B214	Enterobacter	99%	FN	MCOLIBLUE
SAMP_65	B215	Enterobacter	63%	FN	MCOLIBLUE
SAMP_68	B49	Citrobacter	100%	FP	MCOLIBLUE
SAMP_69	B92	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_70	B94	Escherichia/Shigella	100%	FP	MCOLIBLUE
SAMP_72	B97	Cronobacter	100%	FN	MCOLIBLUE

**Table 3.** Sequencing Results for IDEXX Colilert Quanti-Tray®

PCR Number used for Submission	Isolate Number	Organism ID	Percent ID	FP/FN	Media Type
SAMP_188	C216	Escherichia/Shigella	100%	FP	COLILERT
SAMP_190	C221	Escherichia/Shigella	100%	FP	COLILERT
SAMP_192	C223	Enterobacter	91%	FP	COLILERT
SAMP_200	C76	Escherichia/Shigella	99%	FN	COLILERT
SAMP_217	MAR-72	Escherichia/Shigella	100%	FP	COLILERT
SAMP_218	MAR-75	Acinetobacter	100%	FP	COLILERT
SAMP_231	MAR-127	Klebsiella	99%	FN	COLILERT
SAMP_232	MAR-129	Escherichia/Shigella	100%	FN	COLILERT
SAMP_233	MAR-158	Klebsiella	100%	FN	COLILERT
SAMP_234	MAR-159	Klebsiella	100%	FN	COLILERT
SAMP_235	MAR-154	Klebsiella	99%	FP	COLILERT
SAMP_236	MAR-163	Escherichia/Shigella	100%	FN	COLILERT
SAMP_237	MAR-168	Enterobacter	98%	FP	COLILERT
SAMP_239	MAR-90	Escherichia/Shigella	100%	FP	COLILERT
SAMP_250	MAR-105	Escherichia/Shigella	91%	FN	COLILERT
SAMP_272	MAR-74	Enterobacter	90%	FP	COLILERT
SAMP_273	MAR-77	Escherichia/Shigella	100%	FP	COLILERT
SAMP_275	MAR-83	Escherichia/Shigella	100%	FP	COLILERT
SAMP_276	MAR-87	Klebsiella	99%	FN	COLILERT
SAMP_277	MAR-110	Escherichia/Shigella	100%	FP	COLILERT
SAMP_282	MAR-117	Escherichia/Shigella	100%	FP	COLILERT
SAMP_283	MAR-118	Klebsiella	99%	FP	COLILERT
SAMP_286	MAR-155	Enterobacter	89%	FP	COLILERT
SAMP_287	MAR-	Enterobacter	92%	FN	COLILERT

	157				
SAMP_295	MAR-200	Escherichia/Shigella	100%	FP	COLILERT
SAMP_301	C41	Enterobacter	77%	FP	COLILERT
SAMP_310	C50	Escherichia/Shigella	100%	FP	COLILERT
SAMP_311	C60	Pseudomonas	100%	FN	COLILERT
SAMP_321	MAR-81	Pseudomonas	100%	FP	COLILERT
SAMP_324	C66	Klebsiella	100%	FP	COLILERT
SAMP_326	MAR-98	Escherichia/Shigella	97%	FP	COLILERT
SAMP_333	C58	Escherichia/Shigella	100%	FN	COLILERT
SAMP_334	C44	Pseudomonas	100%	FP	COLILERT
SAMP_335	MAR-96	Enterobacter	81%	FP	COLILERT
SAMP_338	C50	Enterobacter	98%	FP	COLILERT
SAMP_339	C46	Escherichia/Shigella	100%	FP	COLILERT
SAMP_73	C123B	Vibrio	100%	FP	COLILERT
SAMP_74	C125B	Shewanella	100%	FP	COLILERT
SAMP_75	C128A	Vibrio	100%	FP	COLILERT
SAMP_76	C129	Vibrio	100%	FP	COLILERT
SAMP_77	C130B	Shewanella	100%	FP	COLILERT
SAMP_78	C135	Shewanella	100%	FN	COLILERT
SAMP_79	C152	Shewanella	100%	FP	COLILERT
SAMP_80	C154	Escherichia/Shigella	100%	FP	COLILERT
SAMP_81	C155	Escherichia/Shigella	100%	FP	COLILERT
SAMP_82	C156	Escherichia/Shigella	97%	FP	COLILERT
SAMP_83	C157	Escherichia/Shigella	100%	FP	COLILERT
SAMP_86	C192	Escherichia/Shigella	100%	FP	COLILERT
SAMP_87	C193	Enterobacter	96%	FP	COLILERT
SAMP_88	C196	Escherichia/Shigella	97%	FP	COLILERT
SAMP_89	C198	Escherichia/Shigella	100%	FP	COLILERT
SAMP_90	C203	Escherichia/Shigella	100%	FP	COLILERT
SAMP_91	C206	Enterobacter	89%	FP	COLILERT
SAMP_94	C45	Enterobacter	98%	FP	COLILERT
SAMP_96	C57	Enterobacter	99%	FN	COLILERT
SAMP_97	C84	Escherichia/Shigella	100%	FP	COLILERT
SAMP_98	C96	Klebsiella	99%	FN	COLILERT

## Appendix D.

### 1) Suggestions of CPS

- a) The research team found the CPS reporting requirements and communication with the grants managers to be pleasant and not burdensome as with some granting agencies. We suggest that CPS strive to maintain this level of commitment to their project PIs in the future.
- b) The research team would like to work with CPS personnel to increase the distribution and dissemination of the research results in the future beyond the term of the grant. This could be in the form of sponsored workshops, conferences, round-table discussions, etc. While the research symposium is a wonderful opportunity to disseminate research results, we would like to build upon this and look towards working with CPS to identify these opportunities in the future and to collaboratively expand upon them.
- c) One suggestion in relation to the comment above is for CPS to host monthly or quarterly “webcasts” where PIs are required to present their final report findings and answer questions to online participants at some point in time during the year following the end of their project. These webcasts can be as short as one hour and have as many participants as needed. On a side note, this is currently required of a grant held by Dr. Rock from the WateReuse Research Foundation. The Foundation presents a webcast each month highlighting the results of new research projects. This mechanism seems to be a successful way to reach/engage a broad number of stakeholders in the industry as well as other research scientists over an extended period of time.

Additionally, Professional Development Hours (PDHs) are offered for these events. A Professional Development Hour is generally defined as one clock hour that is spent engaged in an activity that contributes to the advancement or enhancement of professional skills or scientific knowledge.