Site-Specific Alternative Recreational Criteria Technical Support Materials For Alternative Indicators and Methods

EPA-820-R-14-011
U.S. Environmental Protection Agency
Office of Water
Office of Science and Technology
Health and Ecological Criteria Division

December 2014

Table of Contents

Section 1: Introduction 1
Purpose 1
Overview 2
Methods for Basis of Comparison 4
Relationship to EPA's Alternate Test Procedure (ATP) 5
Determining Whether to Pursue Comparison of Indicator/Methods 6

Section 2: Step-by-step Guide 7
Step 1: Document the Performance of the Alternative Assay 7
Step 2: Gather Water Quality Data 14
Step 3: Compare the Two Indicator/Methods 16

Section 3: Calculate the Site-specific Water Quality Criteria 20
Magnitude 20
Geometric Mean 20
Statistical Threshold Value 22
Duration and Frequency 24

Section 4: WQS Submission Checklist 25

References 26

Appendix A: Factors that Determine Whether to Pursue Comparison of Indicator/Methods A-1

Appendix B: Example Sampling and Analysis Plan B-1

Appendix C: Explanation of Thresholds C-1

Appendix D: Case Examples D-1

Appendix E: How to use Excel to calculate R-squared and index of agreement E-1

Appendix F: Example R Code F-1
Attachment 1: Required Format for Raw Data .csv File G-8
Attachment 2: Required format for comparisons .csv file F-9
Attachment 3: Format of the Output csv File F-10

Tables

Table 1. Four categories of sample results 9

Table 2. Example performance characteristics for EPA Methods 1600 and 1603 13

Table C-1. Data sets used (sorted in descending order of IA values) C-6

Table C-2. Data sets used (sorted in descending order of R-square values) C-10

Table C-3. Methods abbreviations C-14

Table F-1: Example of raw data .csv file (required input for the program) F-8

Table F-2: Example of input comparisons .csv file (required input for the program) F-9

Table F-3: Example of output .csv file (sample output from the program) F-11

Figures

Figure 1. Flow diagram for considering approaches to alternative site-specific criteria 3

Figure 2. Enterococci concentrations for paired samples for enterococci measured by membrane filtration (EPA Method 1600) and Enterolert 18

Figure 3. Comparison of methods for measuring water quality 22

Figure A-1. Idealized relationship between two fecal indicators as the result of their net loading from all sources A-3

Figure C-1. Cumulative distribution of IA values coded by assay-types compared C-3

Figure C-2. Cumulative distribution of R-squared values coded by assay-types compared C-4

Figure D-1. Site A enterococci (culture) and enterococci IMS-ATP data presented graphically D-2

Figure D-2. Site B data presented graphically D-3

Figure D-3. Site C *C. perfringens* and enterococci data presented graphically D-4

Figure E-1. Image of columns A and B (left to right) E-1

Figure E-2. Image of columns C and D (left to right) E-2

Figure E-3. Image of columns E and F (left to right) E-2

Figure E-4. Image of columns G and H (left to right) E-3

Figure E-5. Example of calculations for RSQ and IA E-3

Acronyms

ATCC	American Type Culture Collection
ATP	Alternate Test Procedure
BacHum	*Bacteroides* qPCR
BDL	below detection level
BLOQ	below the limit of quantification
CE	cell equivalents
CCE	calibrator cell equivalent
CFU	colony forming units
CFR	Code of Federal Regulations
CV	coefficient of variation
CWA	Clean Water Act
DNA	deoxyribonucleic acid
ENT-qPCR	*Enterococcus* qPCR
E. coli	*Escherichia coli*
EPA	Environmental Protection Agency
FC	fecal coliform
FIB	fecal indicator bacteria, (e.g., *E. coli*, enterococci)
FN	false negative
FNR	false negative rate
FP	false positive
FPR	false positive rate
GE	genomic equivalents
GM	geometric mean
IA	index of agreement
IDC	initial demonstration of capability
IMS	immunomagnetic separation
IMS-ATP	immunomagnetic separation/adenosine triphosphate
ISO	International Organization for Standardization
LOD	limit of detection
$\log_{10}$	base 10 logarithm
LOQ	limit of quantification
mL	milliliters
MF	membrane filtration
MPN	most probable number
mTEC	membrane-thermotolerant *E. coli* agar
µL	microliters
NEEAR	National Epidemiological and Environmental Assessment of Recreational Water
NTU	nephelometric turbidity units
PCR	polymerase chain reaction
PMA	propidium monoazide
qPCR	quantitative polymerase chain reaction
QA/QC	quality assurance/quality control

QMRA	quantitative microbial risk assessment
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
R-squared	Pearson's correlation coefficient squared
RWQC	recreational water quality criteria
S	precision
SAP	sampling and analysis plan
SCCWRP	Southern California Coastal Water Research Project
SD	standard deviation ($\log_{10}$)
SM	Standard Methods
States	states, tribes, and territories of the United States
STV	statistical threshold value
TC	total coliform
TN	true negative
TP	true positive
TSM	technical support material(s)
U.S.	United States
WQC	water quality criteria
WQS	water quality standard(s)

Section 1: Introduction

Purpose

The United States (U.S.) Environmental Protection Agency (EPA) anticipates that scientific advancements will provide new technologies for quantifying indicators of fecal contamination. These new technologies also might provide improvements over existing methods with regards to rapidity, sensitivity, specificity, and method performance for site-specific applications. As new or alternative indicator and enumeration method combinations[1] mature, states, territories, and authorized Indian tribes (hereafter referred to as states) might want to consider using these methods to develop site-specific water quality criteria (WQC). Information demonstrating that site-specific alternative criteria are scientifically defensible and protective of the recreational use is necessary to support new or revised water quality standards (WQS).

This document provides support materials for developing site-specific alternative WQC using new methods for fecal indicator detection or enumeration that EPA has not validated and issued. This document is part of a set of Technical Support Materials (TSMs) discussed in Section 6 of the 2012 Recreational Water Quality Criteria (RWQC). To best understand this TSM, EPA recommends that you be familiar with the 2012 RWQC and the TSM Guide[2] (U.S. EPA, 2012a; U.S. EPA, 2014). This TSM applies to cases where states wish to use an alternative indicator/method at a site because it has certain advantages over the EPA-recommended methods. This TSM outlines the scientific information needed before an alternative indicator/method can replace the use of a recommended or approved method on a site-specific basis. States may replace the original method at sites where they have demonstrated that an alternative indicator/method has a consistent and predictable relationship with the original method. A state WQS using a different indicator organism or analytical method must be scientifically defensible and protective of the primary contact recreational use. EPA uses WQS for multiple Clean Water Act (CWA) purposes, so in the WQS submission you should discuss the application of the new WQS in the context of specific CWA purposes.

EPA is providing the process outlined in this TSM so that you can use new technology on a site-specific basis. If you answer “yes” to both of the following questions, this TSM might be a useful tool for you to derive site-specific alternative criteria.

---

[1] The term “alternative indicator/method” refers to a method that you would like to use in place of the EPA indicator/method. The alternative indicator/method is also called “method two” in the sections that explain how to conduct the statistical analyses to compare methods. Method two might be “new” in this present application, but might not necessarily be a newly developed method. Method two can be for a different organism, or for an organism that has been used previously for WQS, but with a different assay, or it can be for a different organism and a different assay.

[2] The other TSMs described in the TSM Guide explain how to develop site-specific alternative criteria for alternative health relationships (*Site Specific Alternative Recreational Criteria Technical Support Materials for Alternative Health Relationships*) or non-human fecal sources (*Site-Specific Alternative Recreational Criteria Technical Support Materials for Predominantly Non-Human Fecal Sources*).

1. Do you have another, possibly newer indicator/method that offers advantages compared to the indicator/method that is already in use? Some possible advantages might include more rapid results, lower cost, and ease of use.
2. Are you satisfied with how the current indicator/method measures the health-based goal (either EPA's nationally recommended 2012 RWQC or your state health-based goal)?

If you are not satisfied with how the current indicator/method measures the health-based goal, you might consider using a different TSM. To decide which of the three TSM documents is best for your situation, read the TSM Guide.

Overview

If you have identified an alternative indicator/method and a waterbody as a candidate for site-specific alternative recreational criteria, you can use the information in Appendix A to consider which site features and method factors might be important. The factors in Appendix A can help you decide if a correlation between methods might be likely or unlikely before you undertake water quality sampling. This TSM assumes that you have identified a possible alternative indicator/method already, so it does not include information on how to identify an alternative.

Once you have identified a candidate site and an alternative indicator/method that you think has desirable attributes, the first step is to document the performance of the assay. Performance attributes include precision (repeatability), accuracy, specificity, sensitivity, robustness, and other attributes that could be applicable. If you determine that the assay performance is acceptable for the intended application as discussed in Section 2 of this document, the next step is to gather water quality data at the site using the alternative indicator/method assay along with the chosen EPA-approved method. You can use the water quality data from the site to evaluate whether the alternative indicator/method correlates to one of the EPA-approved indicator/methods. If the association between the two indicator/methods, as measured by an index of agreement (IA) or a Pearson's correlation coefficient squared (R-squared) value, passes specified thresholds, you can derive site-specific water quality criteria. If the new indicator/method does not pass the thresholds for IA or R-squared, you should consider whether you expect the alternative indicator/method to correlate with illness. If you think the alternative indicator/method might correlate with illness, you might want to evaluate this using procedures in the TSM for alternative health relationships (U.S. EPA, 2014). The 2012 RWQC recommendations apply to all waterbodies the state designates for primary contact. If the process in this TSM does not fit your situation, you can choose to retain the EPA recommended criteria and the associated recommended indicator/method, at any point in the process. We describe each step outlined above in more detail in Section 2 of this TSM. Figure 1 is a flow diagram of the process to determine whether an alternative indicator/method is appropriate for your site.

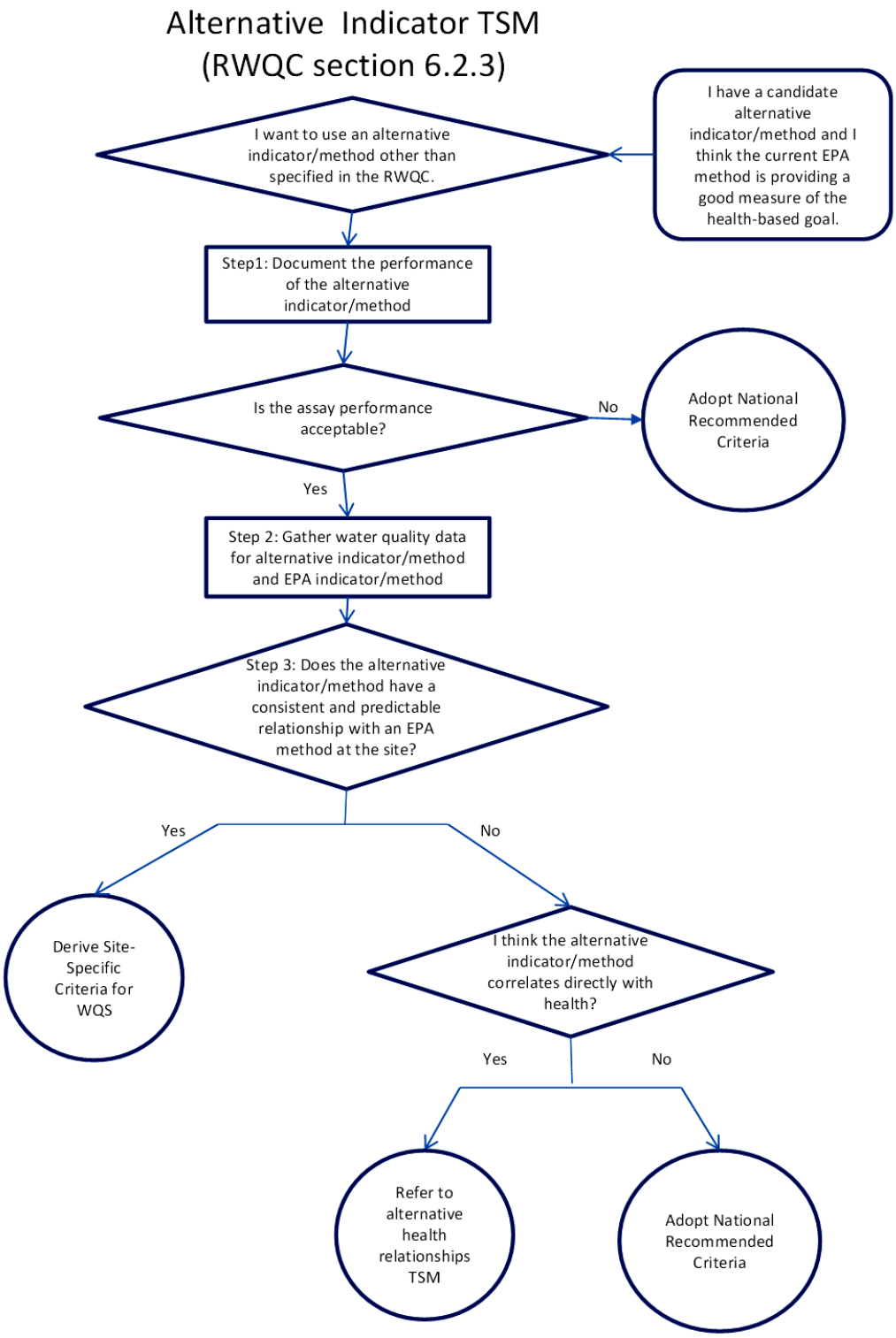

Figure 1. Flow diagram for considering approaches to alternative site-specific criteria

Methods for Basis of Comparison

The approaches described in Section 2 of this TSM compare two analytical methods. Each approach involves an analytical method that is the basis of comparison (method one) and a second analytical method (method two, the alternative indicator/method) that is tested for how well it agrees with method one. In this TSM, method one can be one of the following:

- EPA Method 1600 for enterococci (U.S. EPA, 2009a)[3]
- EPA Method 1603 for *Escherichia coli* (*E. coli*) (U.S. EPA, 2009b)[4]
- EPA Method 1611 for *Enterococcus* spp. (U.S. EPA, 2012b)[5]
- Any "equivalent" method to those listed above, as determined by the Alternate Test Procedure (ATP) program
- Future methods that EPA recommends for CWA §304(a) ambient water quality criteria

When the methods meet the thresholds for the statistical approaches in this TSM, method two can replace method one for ambient water quality monitoring at a specified site. If a state has WQS with fecal coliform (FC) or total coliform (TC), the state may use this TSM, but method one would still need to be one of the methods listed above, not the methods for FC or TC.

After you demonstrate that the alternative indicator/method meets the thresholds for the approach in this TSM, you may adopt it into site-specific WQS (see Section 4). The alternative indicator/method, can be (1) a different fecal indicator organism or substance (e.g., not enterococci or *E. coli*); (2) a different method for the current fecal organisms (e.g., not one of the methods listed above); or (3) a different indicator organism (or substance) with a different method. The alternative indicator/method is not meant to represent a combination of two or more measurements (for example, salinity and a human-specific marker in *Bacteroidales* used in conjunction as a tiered set of indicators).

The approach EPA presents in this TSM indirectly allows for linkage of an alternative indicator/method to human health without the need for conducting any additional epidemiological studies. The human health relationships EPA developed using the National Epidemiological and Environmental Assessment of Recreational Water (NEEAR) studies serve as the basis for 2012 RWQC and can serve as a linkage to an alternative indicator/method.

[3] http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/method_1600.pdf
[4] http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/method_1603.pdf
[5] http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/Method-1611-Enterococci-in-Water-by-TaqMan-Quantitative-Polymerase-Chain-Reaction-qPCR-Assay.pdf

Relationship to EPA's Alternate Test Procedure (ATP)

An ATP is a modification of an EPA-approved reference method that uses the same determinative technique (i.e., the physical and/or chemical process used to determine the identity and concentration of an analyte(s) and measures the same analyte(s) of interest as the EPA-approved reference method. Even if two different analytical methods measure the same analyte(s), EPA considers them to be different methods (U.S. EPA, 2010a). For a new method to become an ATP method, it must undergo the approval process within the ATP program.

Differences between an ATP and an Alternative Indicator/Method Addressed in this TSM

ATP	Alternative Indicator/ Method
• Can be used for national or limited use	• For site-specific use only
• Same analyte only	• Different analytes can be considered

Under the ATP program, an organization or individual may apply for approval of an ATP or new method proposed as an alternative to an EPA-approved reference method.[6] The applicant is responsible for characterizing the performance of the proposed method before submitting it to the ATP program. The Agency reviews the ATP package (which includes comparative data between the proposed method and the EPA-approved reference method) and approves or disapproves the application. For nationwide application, EPA generally will propose to include successful ATPs in the Code of Federal Regulations (CFR), unless the ATP is for limited use or constitutes a minor modification.

The primary intent of a limited-use ATP is to allow use of an ATP or new method by a single laboratory. EPA allows regulated entities to apply limited-use ATPs to one or more matrix types regulated by the CWA or Safe Drinking Water Act (e.g., a specific type of wastewater, ambient water, drinking water). The primary intent of a nationwide-use ATP is to allow use of an ATP or new method by all regulated entities and laboratories for one or more matrix types, including drinking water. Nationwide-use approval allows vendors to establish that new methods produce results that are equivalent to or better than results from methods approved in 40 CFR part 136 (for CWA uses) for compliance monitoring purposes. Nationwide-use approval also allows environmental laboratories across the United States to apply new technologies or modified techniques to more than one matrix type. If a method developer intends to apply the method to more than one matrix type, the developer needs to conduct method studies in each matrix type. This TSM is appropriate only for site-specific (not national) comparisons; furthermore, the alternative indicator/method you might evaluate using this TSM is applicable only for ambient water monitoring (i.e., not wastewater or drinking water).

[6] For details on the ATP program, including the protocol to apply, see: http://water.epa.gov/scitech/methods/cwa/atp/index.cfm. http://water.epa.gov/scitech/methods/cwa/atp/upload/micro_atp_protocol_sept2010.pdf

For the ATP program, the method must measure the same analyte or analytes as an EPA-approved reference method. The provisions in this TSM differ from those in the ATP program in that users can compare a different analyte or analytes (e.g., organism including its enumeration method) to the EPA approved method. The ATP program uses environmental matrices that laboratories spike with the analyte of interest. For this TSM, laboratories analyze environmental samples without spiking. Additional water quality data at a site could in some cases support comparisons between different organisms or other indicator substances (e.g., caffeine, detergent brighteners).

Determining Whether to Pursue Comparison of Indicator/Methods

Several factors that influence whether two microbial methods will yield similar results (i.e., correlate) should be considered before you decide to pursue the approach in this TSM. Later steps in this TSM require water quality monitoring, so before you invest resources in monitoring, you should understand the site-specific factors and assay-related factors that apply to your current and alternative indicator/method. If you decide to proceed, Section 2 outlines the steps in the process. Your site-specific WQS submission should discuss the topics below and any impacts of your results.

Factors that can influence the correlation between indicator/method pairs include:

- Type of assay – The type of assay influences the results. For example, culture methods to enumerate an organism might give different results from molecular methods to enumerate that same organism.
- Fecal sources – The contributions (loadings) of indicator organisms from different fecal sources and the abundance of indicator organisms from all sources, including non-fecal sources, influence the level of fecal indicator bacteria (FIB) in the water body. The predominant source of each type of indicator organism can vary over time.
- Age and proximity of fecal sources – The age and distance the contamination travels can influence the relative abundance of indicator organisms, nonviable chemical indicators, and pathogens with different die-off rates.
- Hydrometeorological[7] factors – Hydrometerological factors influence the loadings from different fecal pollution sources that reach a site and the travel time between the fecal source and receiving water, which in turn determines the age of fecal pollution reaching a site. Loads of indicator organisms may differ between wet weather and dry weather, or between diverse tidal conditions.

Appendix A contains more information on these factors.

[7] Hydrometeorology includes the disciplines of both meteorology and hydrology. It is the study of the transfer of water and energy between the land surface and the lower atmosphere.

Section 2: Step-by-step Guide

This section walks users through three steps for demonstrating the utility of an alternative indicator/method for site-specific WQC. The first step is to document the performance of the new assay. The second step is to gather water quality data at your site using both EPA and alternative methods. The third step is to statistically compare the indicator/method to an EPA method. Figure 1 above shows this step-by-step approach, and a text box summarizes each step below.

Step 1: Document the Performance of the Alternative Assay

New assays might or might not be ready for applications beyond research. Methods that water quality research laboratories find useful might not be appropriate for use in WQS. For example, precision and accuracy may not yet be characterized for these new methods. Documenting the performance of the new method and comparing it to other methods already in use for the same purpose is, therefore, critical. This first step outlines the validation process and presents the performance characteristics for the EPA Methods the 2012 RWQC recommends.

Step 1: Documenting Assay Performance

Information EPA provides

- Documentation for EPA indicator/methods (U.S. EPA, 2009a, 2009b, 2012b)
- List of attributes that should be documented for the alternative indicator/method

Information You Provide

- Information on the performance of the alternative indicator/method based on the list of attributes in this TSM

Decision

- Method performance is understood and deemed acceptable and rationale is provided

The following information is from EPA's *Method Validation of U.S. Environmental Protection Agency Microbiological Methods of Analysis* (U.S. EPA, 2009c). Method validation provides evidence that a specific method can serve its intended purpose. In this case, the evidence would be that the alternative indicator/method detects or quantifies a particular microbe (or group of organisms, or a viral particle) with adequate precision and accuracy. A single laboratory may validate the alternative method, if that laboratory is the only one that will be analyzing samples for the WQS monitoring. If multiple laboratories will evaluate water quality samples, you should arrange for multilaboratory validation.

Tier 1 validation refers to new methods or method modifications that a single laboratory will use with one or more matrix types (i.e., air, water, soil).[8] Validation requirements for Tier 1 reflect this limited use and typically require single-laboratory testing in the matrix types that will use the method. Another name for this type of study is primary validation. Under Tier 1, single laboratories can use methods without having the burden of conducting an interlaboratory method validation study. You should not confuse these studies with laboratory

[8] The matrix type for this TSM is ambient water. U.S. EPA (2009c) also describes Tiers 2 and 3 validations. If you plan to use multiple laboratories for sample analysis, please refer to the Tiers 2 and 3 descriptions.

proficiency testing; proficiency testing is normally associated with evaluating or accrediting laboratories in performing a previously validated and accepted method.

Once an analytical method has been developed and optimized, the first validation step is to determine its operational limits and the within-laboratory performance attributes within these limits that are relevant to the intended use. As with optimization, the laboratory that developed the method often carries out this process, but the organization that intends to implement the method might also complete this first validation step. This primary validation should provide preliminary baseline specifications (numerical and descriptive) of the method's performance within the laboratory performing the tests. The performance attributes that you need to determine can differ depending on the nature and application of the method (e.g., whether the method is culture- vs. microscopy-based, molecular- or chemistry-based, and qualitative vs. quantitative). Guidelines for determining performance attributes of several different broad categories of methods are available (ISO, 1994a; ISO, 1994b; ISO, 2000; U.S. FDA, 2001).

You also need to determine the experimental designs that are best suited to evaluate performance attributes. To obtain examples of relevant experimental designs, you can consult the scientific literature for descriptions of similar, previously validated methods. Note that the same performance attributes might have different terms and definitions in different types and applications of methods. The following sections provide some performance attributes and operational limits that require determination for primary validation of most analytical methods.

Specificity and Sensitivity

Specificity and sensitivity can have different definitions for different types or categories of analytical methods. In a general sense, EPA defines these terms by the extent to which a method responds uniquely to the specified target organism or group of organisms. Specificity is the method's ability to discriminate between the target organism and other (i.e., nontarget) organisms. The mathematical expression for specificity is:

Specificity = TN / (TN + FP)

Where:

TN = Number of samples that correctly tested negative (true negative)
FP = Number of samples that incorrectly tested positive (false positive)

Specificity for microbiology methods and media is traditionally demonstrated by using pure positive and pure negative control cultures. For example, Section 5.1.6.4 of EPA's *Manual for the Certification of Laboratories Analyzing Drinking Water, 5th Edition*[9] (U.S. EPA, 2005) lists the appropriate American Type Culture Collection (ATCC) strains for several groups of enteric control culture bacteria. Positive control cultures listed for enterococci include *Enterococcus faecalis* ATCC 11700 and *Enterococcus faecium* ATCC 6057. Negative controls include *Staphylococcus aureus* ATCC 6538, *E. coli* ATCC 8739 or 25922, and *Serratia marsecens* ATCC

[9] http://water.epa.gov/scitech/drinkingwater/labcert/index.cfm

14756. The specificity is equivalent to 1 minus the false positive rate (FPR), where the FPR is the number of false positives (FP) divided by the total number of samples that are truly negative (FP+TN).

The definition of appropriate target and nontarget control cultures or other standards for use in both validation and routine quality control is an important component of the development of any new microbial method. In a robust method, a single target organism is discernible in complex matrices containing potentially millions of nontarget organisms. Therefore, for the approach in this TSM, primary validation requires demonstration of specificity in environmental samples, in addition to laboratory samples. An independent method should confirm what the method detects in environmental samples. For example, for quantitative polymerase chain reaction (qPCR) methods, during method development the amplified deoxyribonucleic acid (DNA) should be sequenced to confirm the identity of the amplified target.

Sensitivity is the proportion of target organisms that the method can detect. The mathematical expression for sensitivity is:

Sensitivity = TP / (TP + FN)

Where:

TP = Number of samples that correctly tested positive
FN = Number of samples that incorrectly tested negative

Typically, repeated testing of serial dilutions of a “known” spike standard generates the data to calculate sensitivity. The sensitivity is equivalent to 1 minus the false negative rate (FNR) which is the number of false negatives (FN) divided by the total number of samples that are truly positive (FN+TP). FNR is sometimes used in lieu of sensitivity to describe method performance. Table 1 illustrates the values with which to calculate the sensitivity and specificity.

Table 1. Four categories of sample results

	Condition Positive	**Condition Negative**
Test Outcome Positive	True Positive (TP)	False Positive (FP)
Test Outcome Negative	False Negative (FN)	True Negative (TN)

Precision

The International Organization for Standardization (ISO) (1994c) defines precision (S) as the closeness in agreement between independent test results obtained under stipulated conditions and expresses this term as the variance, standard deviation (SD), or coefficient of variation (CV) of a series of test results, where:

%CV = (SD of measurements/mean) × 100

Microbiologists typically derive the expected precision of culture-based microbial methods mathematically, based on the assumption that distribution of bacteria in a well-mixed sample is random and follows a Poisson distribution. Standard Methods (SM) 9222B.6.c (Clesceri et al., 1998) gives the 95% confidence limits around results for methods involving direct counts, such as membrane filtration (MF) or plate counts. Most probable number (MPN) analysis of presence/absence tests of samples divided into multiple tubes or wells provides quantitation. For results obtained by MPN tests, SM 9221C.2 (Clesceri et al., 1998) gives the 95% confidence limits. You should report expected levels of precision for any new proposed method. You can consider precision at four levels: within-laboratory repeatability, within-laboratory reproducibility, between-laboratory repeatability, and between-laboratory reproducibility. You should address the first two levels in the primary validation of a method over the entire density range of the analyte (or microorganism) that you expect to be relevant to its intended use.

For the purpose of this TSM, we define repeatability as the closeness in agreement between results of successive measurements of the same analyte (or microorganism) carried out under the same measurement conditions over a short interval of time. Repeatability is also termed intra-assay precision. Assume that S_r is the within-laboratory precision and S_L is the between-laboratory precision. Then the precision S_R (including within and between) among laboratories is expressed as:

$$S_R = \sqrt{S_r^2 + S_L^2}$$

For this TSM, we define reproducibility as the closeness of the agreement between the results of measurements carried out on the same analyte under variable conditions of measurement. For determination of within-laboratory reproducibility, some of the variable conditions you should consider include different time intervals between analyses, more than one analyst, numerous preparations of reagents, different instruments, and different water matrices.

Accuracy and Bias

One definition of accuracy is the closeness of the agreement between a test result and the accepted reference value. A definition of bias is the difference between the expectation of the test results and a known or accepted reference value. ISO (1994c) defines the term “accuracy,” when applied to a set of test results, more comprehensively as a combination of random components (related to random error) and a common bias component (related to total systematic error) associated with the method. Like precision, you also should first characterize this bias component at the within-laboratory level as a primary attribute of most methods. Given the above definition, you determine bias in the analysis of a material by a new method by knowing the true value for the analyte in the sample or assigning an accepted reference value. In some cases, using fortified or spiked samples, certified reference materials, analysis by another presumably unbiased method, or internal controls will help you determine or assign these values. When analyzing fortified or spiked samples, researchers often express bias in terms of the recovery or percent recovery, that is, the test result, divided by the expected (assigned) value for the added spike material, multiplied by 100.

In some instances, a direct means of determining the bias of a new method's test results might not be available. Under these circumstances, you can sometimes assess the recovery of the analyte by the new method in relation to the results of an accepted reference method and express it as "relative recovery" (ISO, 2000). If no such reference method is available, you can define relative recovery by the new method itself. To define relative recovery by the new method, you should compare the test results from two simultaneously processed and analyzed samples with unknown quantities of analyte. In this case, you can express the two test results as the ratio of the analyte recovered in the two samples. By designating one of these unknown samples as a reference, you can determine the relative recovery of the analyte from any number of additional, simultaneously processed and analyzed unknown samples and compare them to each other based on their respective recovery ratios. Even for methods with high variability, this technique is useful because you can quantify the variation by measuring SD. In addition, if you can consider the reference sample to be consistent material having the same (albeit unknown) quantity of analyte in different test runs of the method, you can also use the respective recovery ratios of this sample to compare the relative recoveries from other unknown samples.

Limit of Quantification (LOQ)

In all methods, technological aspects limit how many organisms the assay needs to return a "positive" result. The limit of detection (LOD) for culture-based enumeration methods is the lowest number of microorganisms distinguishable from the absence of microorganisms. The LOD is usually one colony per plate (or membrane). How the water samples are concentrated or diluted determines the volume of water associated with each plate, so the LOD can vary from assay to assay and laboratory to laboratory. The limit of quantification (LOQ) is the lowest quantity that an assay can reliably enumerate. For example, EPA Method 1600 assumes that the acceptable range of counts is between 20 and 60 colonies per membrane (U.S. EPA, 2009a). The LOD is one colony per plate, but the LOQ is 20 colonies per plate. The data that EPA used in this TSM was adjusted by the contributing researchers for sample volume. Therefore we could not determine whether the values reported were close to the LOD or LOQ of the assay. Many laboratories treat one colony on a plate as a quantifiable count, so culture-based methods often treat LOD as the LOQ.

For molecular methods, the LOQ and LOD are clearly different. qPCR can detect small quantities of DNA target sequences, but cannot reliably quantify small quantities. For example, if one DNA molecule is detected in a reaction that represents 1 milliliter (mL) of a 100 mL original water sample, the LOD would be 100 polymerase chain reaction (PCR) units per 100 mL.[10] The LOQ might be three times higher (300 PCR units per 100 mL), but it might also be 10 times higher (1,000 PCR units per 100 mL). You should characterize and document the LOD and LOQ of the alternative indicator/method.

[10] 100 mL of water can be reduced to 100 microliters (µL) of DNA template. One reaction might use 1 µL of DNA template. In addition there are a variety of ways to translate PCR units to numbers of organisms or numbers of genomes.

Documentation of the Method

The next step in the validation process is to prepare a complete written description of the method. Historically, EPA has written its methods in a format that includes the following components:

- scope and application;
- method summary;
- definitions;
- interferences;
- health and safety;
- equipment and supplies;
- reagents and standards;
- sample collection, preservation, and storage;
- quality control;
- calibration and standardization;
- procedural steps;
- calculations and data analysis;
- method performance;
- pollution prevention; and
- waste management.

Appendix A of *Method Validation of U.S. Environmental Protection Agency Microbiological Methods of Analysis* (U.S. EPA, 2009c) describes each component of this format in detail. Note that EPA recommends this format, but it is not required. You should indicate in the method documentation whether your methodology can be implemented with next generation equipment and reagents or other vendor equipment. For example, you should ensure that references to specific brands or catalog numbers in the method do not preclude the use of other vendors, equipment, or supplies.

Your method description must address quality assurance and quality control (QA/QC). EPA Method 1601 (U.S. EPA, 2001a) presents an example of a QA/QC description for a method.[11] Quality control includes an initial demonstration of capability (IDC). You should perform IDC to demonstrate acceptable performance with the method as written before you analyze field samples or evaluate acceptable performance of a method modification. In addition to IDC, you should also describe ongoing demonstration of capability documentation in the QA/QC section of the method.

If your new indicator/method is a qPCR or digital PCR method, then publications on the minimal information required for these types of analyses (Bustin et al., 2009; Huggett et al., 2013) should be referenced when describing details of the assays.

[11] http://www.epa.gov/microbes/documents/1601ap01.pdf

Example Performance Characteristics

Table 2 documents some of the performance characteristics for the recommended EPA indicators/methods to help you determine whether the thresholds for your performance characteristics are reasonable given what other methods can achieve.

Table 2. Example performance characteristics for EPA Methods 1600 and 1603

Performance Characteristic	EPA Method 1600	EPA Method 1603
False positive rate	6.0%	6%
False negative rate	6.5%	5%
Precision	2.2% (for marine water) 18.9% (for surface water)	Not quantified[12]

As Table 2 shows, the microbiological medium used in EPA Method 1600 has a 6.0% false positive rate (94% specificity) and 6.5% false negative rate (93.5% sensitivity) for various environmental water samples (U.S. EPA, 2009a). Regarding bias, the persistent positive or negative deviation of the results from the assumed or accepted true value was not significant. The precision among laboratories for marine and surface water was 2.2% and 18.9%, respectively (U.S. EPA, 2009a). For EPA Method 1603, the false-positive and false-negative rates were 6% and 5%, respectively (U.S. EPA, 2009b). For examples of how to document performance characteristics, refer to EPA Methods 1600 and 1603, Section 15.0 Method Performance.

Whether a validated method is successful in practice can depend on many factors. For example, within ISO, at least 75% of the member bodies casting a vote must approve the acceptance and publication of a method as a standard method (WHO, 2003).

Once you have documented the performance of the method and shown that the specificity, sensitivity, precision, repeatability, reproducibility, accuracy, bias, and limits of quantification are similar to what EPA published methods have achieved, you can proceed to Step 2 and continue with the evaluation process.

[12] Method 1603 says "Precision – The degree of agreement of repeated measurements of the same parameter expressed quantitatively as the SD or as the 95% confidence limits of the mean computed from the results of a series of controlled determinations. The modified membrane-thermotolerant *E. coli* agar (mTEC) method precision was found to be fairly representative of what would be expected from counts with a Poisson distribution. Bias – The persistent positive or negative deviation of the average value of the method from the assumed or accepted true value. The bias of the modified mTEC method has been reported to be -2% of the true value."

Step 2: Gather Water Quality Data

Step 2: Gather Information and Water Quality Data

Information EPA provides

- Example sampling and analysis plan (Appendix C)

Information You Provide

- Sampling and analysis plan
- Water quality data

Decisions

- Data are adequate. For example, 30 paired data points are within the limits of quantification.

To compare your alternative indicator/method to an established method, you need water quality data measuring the densities of organisms with both methods. You should collect data over the year or the recreational season to reflect the range of environmental conditions, such as meteorological conditions (for example, rain events and dry weather events) and hydrodynamic changes (due to winds, tides, or other extreme events frequently encountered at the site). It may be important to collect samples during and after phytoplankton blooms and during and after high turbidity events if these conditions typically occur at the field site (Wymer et al., 2005). In some areas, the recreational season can be as short as 3 months, while in others it might be the entire year. You should include all relevant types of conditions (wet, dry, tidal, bloom, turbid) during sampling.

To be confident in the analysis that you will complete in Step 3, you need at least 30 paired data points (within the quantification limits of the assays) in the data set. You should obtain these paired data from 30 samples taken at intervals over time to cover the range of conditions at the site. The choice of at least 30 samples within the limit of quantification has precedent in the 2004 Beach Rule (U.S. EPA, 2004). However, it may be desirable and necessary to include more samples than 30 in the analysis. If you find that it is difficult to collect 30 samples above the limit of quantification of the two assays, then you might wish to utilize the companion TSM, *Site Specific Alternative Recreational Criteria Technical Support Materials for Alternative Health Relationships.*

Clear documentation of the sampling plan is essential. You should include the sampling and analysis plan (SAP) in your WQS submission to provide supporting documentation for a site-specific alternative WQC. You should include sufficient detail in the SAP to allow review of the study design as to whether it is scientifically defensible. In brief, an SAP would include

- a map of the site with sampling locations labeled;
- water depth at sampling locations;
- time of day of sampling;
- other information that will be collected (such as weather, sunlight, bird counts, or other parameters);
- frequency of sampling;
- sample holding times;
- enumeration methods (including limits of quantification); and
- explanation of how the environmental sampling is representative of conditions expected at your site.

The SAP also should include how you will conduct QA/QC. Appendix B presents an example SAP.

Step 3: Compare the Two Indicator/Methods

In Step 3, you compare the alternative indicator/method (method 2) to the established indicator/method (method 1) using analysis of the paired water quality samples gathered in Step 2. You perform these analyses to demonstrate a consistent and predictable relationship between method 1 and 2, which has a health relationship informed by epidemiological studies (U.S. EPA, 2012a). Step 3 describes how to prepare the raw water quality data for analysis and how to calculate an index of agreement (IA) and the R-squared value.

Step 3: Compare Indicators/Methods

Information EPA provides

- Statistical methods (IA and R-squared) for comparing indicator methods
- Example case studies for illustrating statistical methods and applying IA and R-squared thresholds, including example spreadsheet

Information You Provide

- Water quality data for the alternative indicator/method and the EPA-indicator/method (or equivalent) for your waterbody
- Transparent documentation of the application of the statistical methods described in this TSM

Decisions and Rationale

- Are the IA or R-squared at or above the thresholds?
- The rationale for your decision.

All waterbodies have some variability associated with water quality. You can describe water quality (as represented by indicator organism density) by a geometric mean (GM) and SD. You must include both the GM and the SD in the analysis documentation.

In Step 1, you documented the LOQ for the alternative indicator/method and the method you are currently using. Use the data handling steps below to prepare the data set collected in Step 2.

1. Omit the paired data points where one or both measurements are below the LOQ.
2. If any data points are above the maximum level of detection, you should remove those paired data points from the data set or dilute the samples and reanalyze them. You also can remove any data from the data set that indicate inhibition or interference with the assay method. You should show the raw data and the data that have been through these treatments.
3. Identifying outliers in environmental data sets can be challenging as it is difficult to discern outliers that are due to measurement errors and anomalies, and stochastic, yet realistic environmental variability. You should eliminate outliers that you consider to be a result of measurement errors or anomalies, but not those associated with environmental variability. A sound justification should be provided for removal of such data points. Outliers that are due to sporadic conditions, such as storm events, are part of the environmental variability and can be included in the data set.
4. Compute the base 10 logarithm (log10) of each of the data points. This computation will transform the data set so that you can approximate the distribution of indicator organism densities by a normal distribution (Wymer and Wade, 2007).

Once you have prepared the data set by applying the four data handling steps above, you can calculate the IA using Equation 1. You can use IA to detect additive and proportional differences between two data sets. Appendix C further explains why EPA chose a statistical test based on IA.

$$IA = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2}$$ [Equation 1][13]

In Equation 1, x represents density of the microorganism determined by Method 1 (EPA Method 1600, 1603 or 1611 or equivalent), and y represents the density of the microorganism determined by Method 2. Also in the equation, *i* is a counter, *N* (sample size) is the total number of data points in the data set, $\bar{x}$ and $\bar{y}$ are the averages of the x and y data sets, respectively, and *IA* varies from 0 to 1 (Willmott and Wicks, 1980).

This TSM document describes two steps to assess the agreement between the two indicator/methods:

Statistical Thresholds for Agreement

IA ≥ 0.7
R-squared > 0.6

1. Calculate IA to assess whether agreement between the two methods is sufficient (that is IA ≥ 0.7). If IA is greater than or equal to 0.7, you can use the alternative indicator/method and the unchanged numerical criteria values for the EPA indicator/method (see Section 3). Data sets with an IA ≥ 0.7 have acceptable agreement. EPA selected this threshold based on the process outlined in Appendix C.
2. If IA does not indicate good agreement between the methods (that is, IA < 0.7), calculate R-squared between the two methods to determine if the indicators are correlated.[14] R-squared is the proportion of variance in Y that you can account for by knowing X. If the R-squared value is greater than 0.6, you may use the alternative indicator/method to derive site-specific alternative WQC, but you need to derive new numerical limits (see Section 3). Appendix E provides instructions on how to calculate R-squared and IA in Excel and Appendix F provides instructions on how to calculate R-squared and IA using the R computational language.[15]

Figure 2 illustrates a data set with good agreement between two indicator/methods. The data in Figure 2 are from Doheny Beach in Dana Point, California. In 2008, Southern California Coastal Water Research Project (SCCWRP) collected water samples and evaluated them using both Enterolert and EPA Method 1600. The IA for this data set is 0.94, which is above the threshold of 0.7. The R-squared for this data set is 0.79, which also is above the threshold of 0.6. Figure 2 has dotted lines at 35 colony forming units (CFU) or MPN per 100 mL, which corresponds to one of the recommended RWQC GM magnitudes in the 2012 RWQC.

[13] 1/N in the numerator and denominator are not included in Willmott and Wicks (1980). This term was added to facilitate the ease of implementing the equation in the software.
[14] You can calculate Pearson's correlation using the 'correl' function in Excel.
[15] R is a free software environment for statistical computing and graphics and is available for download at http://www.r-project.org/

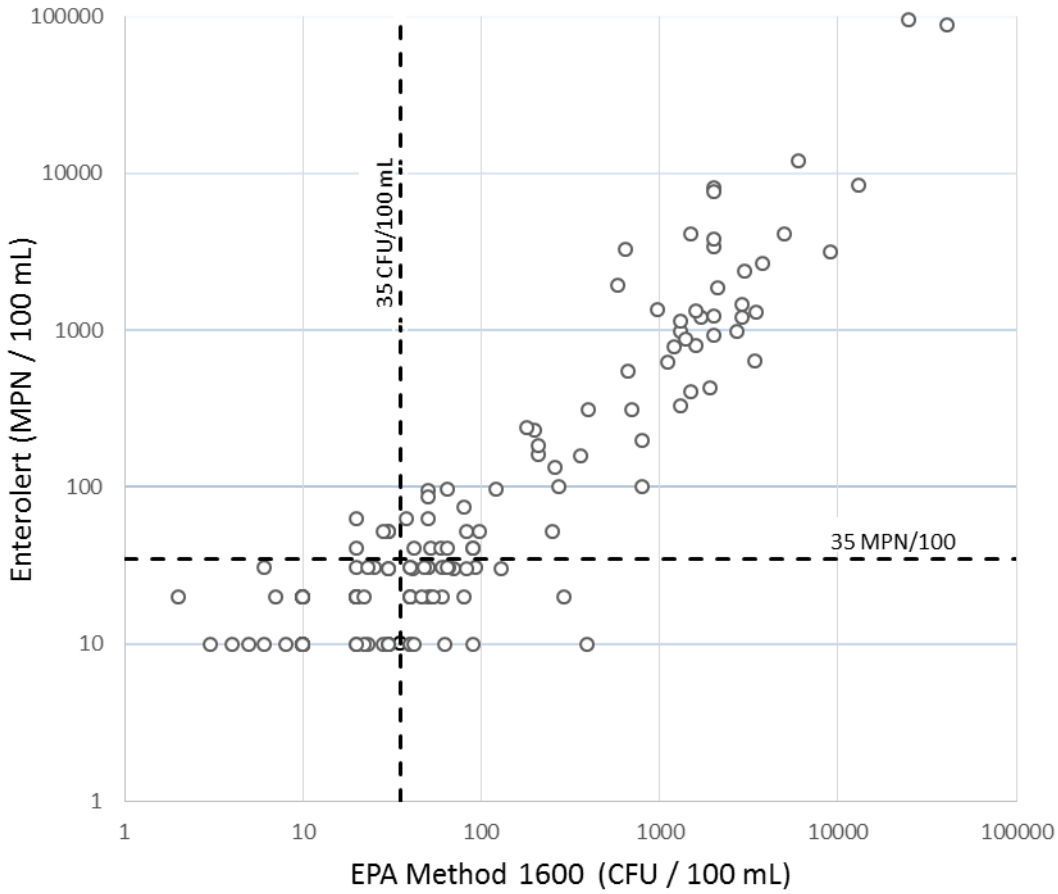

Figure 2. Enterococci concentrations for paired samples for enterococci measured by membrane filtration (EPA Method 1600) and Enterolert

Data were $\log_{10}$ transformed and samples below the LOQ were omitted (n above LOQ = 148). SCCWRP contributed these data, which are from Doheny Beach in Dana Point, California, collected in 2008.

Appendix C explains the selection of the IA and R-squared thresholds. Appendix D contains three examples to illustrate how the comparison of two indicator/methods is conducted using the statistical methods in this TSM.

If your indicator/method comparisons do not meet the IA or the R-squared thresholds, you cannot use this TSM to derive site-specific alternative WQC. You may still, however, evaluate your alternative indicator/method using the TSM *Site-Specific Alternative Recreational Criteria Technical Support Materials for Alternative Health Relationships.*

Over time, conditions that influence FIB dynamics can change, such as land use patterns. You should, therefore, reevaluate your site-specific alternative criteria every three years,[16] as part

[16] Under EPA's WQS implementing regulations at 40 CFR 131.20(a), states must hold public hearings at least once every 3 years to review applicable WQS, and, as appropriate, modify and adopt standards.

of your state's triennial revision of WQS. The reevaluation is needed to confirm that the relationship between the indicator/methods has remained valid.

Section 3: Calculate the Site-specific Water Quality Criteria

RWQC include a magnitude, duration, and frequency. Magnitude is the numeric expression of the maximum amount of the pollutant that may be present in a waterbody that supports the designated use. Duration is the period of time over which the magnitude is calculated. Frequency of excursion describes the maximum number of times the pollutant may be present above the magnitude over the specified period (duration). A WQS consists of a magnitude, duration, and frequency, must be scientifically defensible, and protect the designated use, in this case, primary contact recreation.

Magnitude

The 2012 RWQC, recommend expressing magnitude as a GM value corresponding to the 50th percentile and a statistical threshold value (STV) corresponding to the 90th percentile of the same water quality distribution; thus both values used together would be associated with the same level of public health protection. EPA's criteria recommendations are for GM and STV (rather than just for GM or just for STV) because, used together, they indicate whether the water quality is protective of the designated use of primary contact recreation. Using GM alone would not reflect spikes in water quality because GM alone is not sensitive to spikes.

Geometric Mean

Using the information from Steps 2 and 3 (in Section 2), you can derive site-specific alternative criteria (Figure 1).

If IA is greater than or equal to 0.7, you may use the alternative indicator/method and your site-specific criteria values would be the same as the numerical criteria values for the EPA indicator/method (see Section 2). For example if the EPA indicator/method is EPA Method 1600, which has a GM of 35 CFU per 100 mL, and IA ≥ 0.7, the alternative indicator/method GM also can be 35 units per 100 mL. If IA ≥ 0.7, you need conduct no further statistical analyses because your data set has acceptable agreement with the EPA method.

If IA is less than 0.7, you can check the comparison of the methods using R-squared as the statistical metric, as described in Section 2, Step 3. Appendix C presents more information on IA and R-squared. If the alternative indicator/method correlates with the EPA method, based on the R-squared value (R-squared > 0.6), you can derive new criteria values through linear regression of the log-transformed data. This approach develops a quantitative relationship between the alternative indicator/method and the EPA indicator/method, which you then can use to derive the new criteria values. The results of the analysis and recommended new criteria value will be reviewed on a case-by-case basis in the context of the standards package submission. In this case, the linear regression determines a line by the statistical method of least squares that best fits the data. The relationship will follow the following formula (refer to Figure 3):

$$Y = mX + b$$

Where:

$Y = \log_{10}$ indicator organism density of the alternative indicator/method

m = slope of the line

$X = \log_{10}$ indicator organism density of the EPA indicator/method

b = y-intercept

You can compute the linear regression using the 'linest' function in Excel. Alternatively, you may use the 'slope' and 'intercept' functions. Once you determine the relationship between the alternative indicator/method and the EPA indicator/method (by calculating the slope and intercept above), you can calculate the new GM value as follows:

$$Y_{Criterion} = 10^{(m * \log_{10}(X_{Criterion}) + b)}$$

Where $Y_{Criterion}$ is the new criterion value and $X_{Criterion}$ is the 2012 RWQC value (or supplemental element) for marine and freshwater:[17]

Marine water GM

- 35 CFU/100 mL or 30 CFU/100 mL for enterococci
- 470 calibrator cell equivalents (CCE)/100 mL or 300 CCE/100 mL for qPCR *Enterococcus*

Freshwater GM

- 126 CFU/100 mL or 100 CFU/100 mL for *E. coli*
- 35 CFU/100 mL or 30 CFU/100 mL for enterococci
- 470 CCE/100 mL or 300 CCE/100 mL for qPCR *Enterococcus*

Figure 3 shows a graphical representation of the above equation.

[17] Note that the 2012 RWQC recommends two possible criteria values for each indicator corresponding with two illness rates. The qPCR values are considered "supplemental elements," whereas the CFU values are the recommended 2012 RWQC. Both the RWQC values and the supplemental elements are shown, because both are acceptable as Method 1.

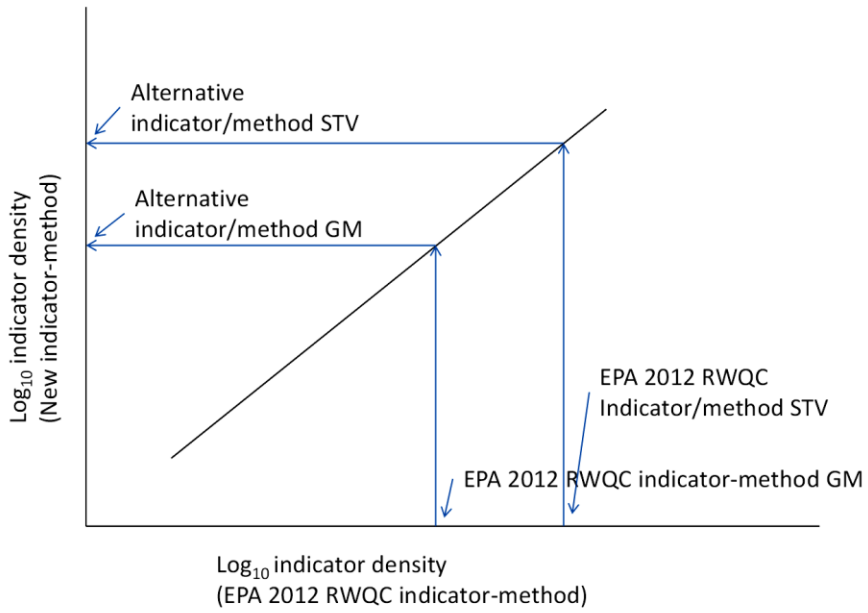

Figure 3. Comparison of methods for measuring water quality

The graph shows how you can translate the GM and STV values for the EPA indicator/method to a GM and STV for the alternative indicator/method, using the relationship between the two indicator/methods that you determine using paired environmental samples.

Statistical Threshold Value

The 2012 RWQC express magnitude as a GM and an STV. EPA bases the 2012 RWQC STV on the observed SD of FIB densities during the NEEAR study.[18] A larger SD reflects a broader range of FIB densities at a site and can occur because the proportion of "high-density" samples in the data set is larger. Days with high densities of FIB correspond to days of higher predicted illness levels. If you use a site-specific SD, the potential exists for a higher proportion of days where the expected illness level would be higher. For example, if a site is subject to combined sewer overflows, the SD could be greater than reported during the NEEAR study. A larger SD results in a higher STV and more days could be below the STV. To be consistent with the level of water quality that EPA considers protective of the recreational designated use, the SD must be equal to or lower than the NEEAR SD. From the NEEAR study sites, the pooled estimate for the log SD of culturable enterococci was 0.44, and EPA previously reported the pooled log SD for

[18] EPA conducted epidemiological investigations at U.S. beaches in 2003, 2004, 2005, 2007, and 2009, and refers to these investigations as a group as the NEEAR study. The NEEAR study enrolled 54,250 participants, encompassed 9 locations, and collected and analyzed numerous samples from fresh water, marine, tropical, and temperate beaches (U.S. EPA, 2010b; Wade et al., 2008, 2010).

culturable *E. coli* is 0.40 (U.S. EPA, 1986; U.S. EPA, 2012a).[19] For *Enterococcus* measured by qPCR, the SD from the NEEAR study is 0.49 (U.S. EPA, 2012a).

If IA is greater than or equal to 0.7, you can use the alternative indicator/method and the site-specific STV for the alternative indicator/method would be the same as the numerical criteria values for the EPA indicator/method (see Section 2). For example, if the EPA indicator/method is EPA Method 1600, for which STV is 130 CFU per 100 mL and IA ≥ 0.7, the alternative indicator/method STV also can be 130 units per 100 mL. Data sets with IA ≥ 0.7 have acceptable agreement. If IA ≥ 0.7, you need conduct no further statistical analyses.

If IA is less than 0.7, you can check the comparison of the methods using R-squared as the statistical metric, as described in Section 2, Step 3. For more information on the IA and R-squared values, refer to Appendix C. If the two indicator/methods are correlated using R-squared (i.e., R-squared > 0.6), you may derive new criteria values through linear regression of the log-transformed data, as described above in deriving the magnitude of the GM standard. In this case, the linear regression determines a line by the statistical method of least squares that best fits the data. The relationship will follow the following formula (refer to Figure 3):

$$Y = mX + b$$

Where:

$Y = \log_{10}$ indicator organism density of the alternative indicator/method

m = slope of the line

$X = \log_{10}$ indicator organism density of the EPA indicator/method

b = y-intercept

You can compute the linear regression with the 'linest' function in Excel. Alternatively, you may use the 'slope' and 'intercept' functions. Once you have determined the relationship between the alternative indicator/method and the EPA indicator/method (by calculating the slope and intercept above), you can calculate the new STV value as follows:

$$Y_{STV} = 10^{(m * \log_{10}(X_{STV}) + b)}$$

Where Y_{STV} is the new STV value and X_{STV} is the 2012 RWQC STV value for marine and freshwater:[20]

Marine water STV

- 130 CFU/100 mL or 110 CFU/100 mL for enterococci

[19] RWQC, page 40.

[20] Note that the 2012 RWQC recommends two possible criteria values for each indicator corresponding with two illness rates. The qPCR values are considered "supplemental elements," whereas the CFU values are the recommended 2012 RWQC. Both the RWQC values and the supplemental elements are shown, because both are acceptable as Method 1.

- 2,000 CCE/100 mL or 1,280 CCE/100 mL for qPCR *Enterococcus*

Freshwater STV

- 410 CFU/100 mL or 320 CFU/100 mL for *E. coli*
- 130 CFU/100 mL or 110 CFU/100 mL for enterococci
- 2,000 CCE/100 mL or 1,280 CCE/100 mL for qPCR *Enterococcus*

Duration and Frequency

The 2012 RWQC recommend that the waterbody GM not be greater than the selected GM magnitude in any 30-day interval (duration). The RWQC also recommend that the excursion frequency of the selected STV magnitude should not be greater than 10% in the same 30-day interval. You should use the same duration and frequency for your site-specific alternative criteria as the 2012 RWQC recommend.

If your indicator/method choice does not pass either the IA or R-squared thresholds, you can still consider another route for deriving site-specific alternative criteria. One option is to develop an alternative health relationship for your alternative indicator/method (see *Site Specific Alternative Recreational Criteria Technical Support Materials for Alternative Health Relationships)*.

Section 4: WQS Submission Checklist

This section provides a checklist of additional information you should include in your WQS submission that includes your alternative site-specific WQC. This list is in addition to the information in EPA's *Water Quality Standards Handbook Chapter 6: Procedures for Review and Revision of Water Quality Standards* (40 CFR 131 Subpart C).[21] The process for submittal and approval of WQS is as indicated in EPA's *Water Quality Standards Handbook*.

Information You Provide

_____Information on the performance of the new method including:

- _____specificity
- _____sensitivity
- _____precision
- _____repeatability
- _____reproducibility
- _____accuracy
- _____bias
- _____limits of quantification

_____Sampling and analysis plan

_____Water quality data for your alternative indicator/method and the EPA indicator/method (or equivalent) for your waterbody

_____Transparent application of the statistical methods in the TSM (Step 3)

_____Transparent application of the approach for deriving site-specific alternative criteria (Section 3)

Decision to Be Captured

_____Method performance is understood and deemed acceptable (Step 1).

_____Data collected in Step 2 are adequate for use in Step 3 (30 paired data points are within the limits of quantification)

_____Is IA or R-squared above the thresholds (Step 3)? (One of the following applies.)

- _____If IA is above the threshold (≥ 0.7), GM and STV are the same numeric value as the EPA indicator/method (method one in the statistical analysis).
- _____If IA is below the threshold and R-squared is above the threshold (> 0.6), the GM value is calculated using the regression equation show in Section 3.
- _____If IA and R-squared are both below the thresholds, this TSM approach cannot be used to derive site-specific alternative criteria.

[21] http://water.epa.gov/scitech/swguidance/standards/handbook/index.cfm
http://water.epa.gov/scitech/swguidance/standards/handbook/chapter06.cfm

References

Agidi, S., Vedachalam, S., Mancl, K., Lee, J. 2013. Effectiveness of Onsite Wastewater Reuse System in Reducing Bacterial Contaminants Measured with Human-specific IMS/ATP and qPCR. Journal of Environmental Management 115: 167–174.

Alm, E.W., Burke, J., Burke, Hagan, E. 2006. Persistence and Potential Growth of the Fecal Indicator Bacteria, *Escherichia coli*, in Shoreline Sand at Lake Huron. Journal of Great Lakes Research 32(2): 401–405.

Badgley, B.D., Nayak, B.S., Harwood, V.J. 2010. The Importance of Sediment and Submerged Aquatic Vegetation as Potential Habitats for Persistent Strains of Enterococci in a Subtropical Watershed. Water Research 44(20): 5857–5866.

Bae, S., Wuertz, S. 2009. Rapid Decay of Host-specific Fecal Bacteroidales Cells in Seawater as Measured by Quantitative PCR with Propidium Monoazide. Water Research 43: 4850–4859.

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., Sayler, G.S. 2009. Factors Influencing the Persistence of Fecal *Bacteroides* in Stream Water. Journal of Environmental Quality 38: 1224–1232.

Boehm, A.B., Grant, S.B., Kim, J.-H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., Wellman, D.E. 2002. Decadal and Shorter Period Variability of Surf Zone Water Quality at Huntington Beach, California. Environmental Science & Technology 36: 3885–3892.

Boehm, A.B., Weisberg, S.B. 2005. Tidal Forcing of Enterococci at Marine Recreational Beaches at Fortnightly and Semidiurnal Frequencies. Environmental Science & Technology 39: 5575–5583.

Boehm, A.B., Yamahara, K.M., Love, D.C., Peterson, B.M., McNeill, K., Nelson, K.L. 2009. Covariation and Photoinactivation of Traditional and Novel Indicator Organisms and Human Viruses at a Sewage-impacted Marine Beach. Environmental Science and Technology 43: 8046–5052.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T. 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical Chemistry 55(4):611–622.

Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Sadowsky, M.J., Ishii, S. 2006. Population Structure, Persistence, and Seasonality of Autochthonous *Escherichia coli* in Temperate, Coastal Forest Soil from a Great Lakes Watershed. Environmental Microbiology 8(3): 504–513.

Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Nevers, M.B. 2010. Linking Non-culturable (qPCR) and Culturable Enterococci Densities with Hydrometeorological Conditions. Science of the Total Environment 408: 3096–3101.

Byappanahalli, M.N., Roll, B. M., Fujioka, R.S. 2011. Evidence for Occurrence, Persistence, and Growth Potential of *Escherichia coli* and Enterococci in Hawaii's Soil Environments. Microbes and Environments 27: 164–170.

Carrillo, M., Estrada, E., Hazen, T.C. 1985. Survival and Enumeration of the Fecal Indicators *Bifidobacterium adolescentis* and *Escherichia coli* in a Tropical Rain Forest Watershed. Applied and Environmental Microbiology 50(2): 468–476.

Chern, E. C., Siefring, S., Paar, J., Doolittle, M., Haugland, R. 2011. Comparison of Quantitative PCR Assays for *Escherichia coli* Targeting Ribosomal RNA and Single Copy Genes. Letters in Applied Microbiology 52: 298–306.

Clesceri, L.S., Eaton, A.D., Greenberg, A.E. 1998. Standard Method 9221F Multiple-tube Fermentation Technique for Members of the Coliform Group (20th Edition) in Standard Methods for the Examination of Water and Wastewater (20th edition). Standard Methods is a joint publication of the American Public Health Association (APHA), the American Water Works Association (AWWA), and the Water Environment Federation (WEF). EC-MUG (Method 9221 F) or NA-MUG (Method 9222 G) can be used for *E. coli* testing step as described in 141.21(f)(6)(i) or (iii) after use of Standard Methods 9221 B, 9221 D, 9222 B, or 9222 C. Available at: http://www.standardmethods.org/store/ProductView.cfm?ProductID=344

Converse, R.R., Blackwood, A.D., Kris, M., Griffith, J.F., Noble, R.T. 2009. Rapid QPCR-based Assay for Fecal *Bacteroides* spp. as a Tool for Assessing Fecal Contamination in Recreational Waters. Water Research 43: 4828–4837.

Converse, R.R., Griffith, J.F., Noble,R.T., Haugland, R.A., Schiff, K.C., Weisberg, S.B. 2012. Correlation between Quantitative PCR and Culture-Based Methods for Measuring *Enterococcus* spp. over Various Temporal Scales at Three California Marine. Applied and Environmental Microbiology 78(4): 1237–1242.

Davies, C.M., Long, J.A., Donald, M., Ashbolt, N.J. 1995. Survival of Fecal Microorganisms in Marine and Freshwater Sediments. Applied and Environmental Microbiology 61(5): 1888–1896.

Dickenson, J.A. and Sansalone, J.J. 2011. Distribution and Disinfection of Bacterial Loadings Associated with Particulate Matter Fractions Transported in Urban Wet Weather Flows. Water Research 46(20): 6704–6714.

Fujioka, R.S. and Byappanahalli, M.N. 2003. Proceedings and Report of Tropical Water Quality Indicator Workshop. University of Hawaii at Manoa. March, 2001. Special Report SR-2004-01.

Grant, S.B., Sanders, B.F., Boehm, A.B., Redman, J.A., Kim, J.H., Mrše, Chu, A.K., Gouldin, M., McGee, C.D., Gardiner, N.A., Jones, B.H., Svejkovsky, J., Leipzig, G.V., Brown, A. 2001. Generation of Enterococci Bacteria in a Coastal Saltwater Marsh and its Impact on Surf Zone Water Quality. Environmental Science & Technology 35(12): 2407–2416.

Hardina, C.M. and Fujioka, R.S. 1991. Soil: The Environmental Source of *Escherichia coli* and Enterococci in Hawaii's Streams. Environmental Toxicology Water Quality 6: 185–195.

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., Dufour, A.P. 2005. Comparison of *Enterococcus* Measurements in Freshwater at Two Recreational Beaches by Quantitative Polymerase Chain Reaction and Membrane Filter Culture Analysis. Water Research 39(4): 559–568.

Huggett, J.F., Foy, C.A., Benes, V., Emslie, K., Garson, J.A., Haynes, R., Hellemans, J., Kubista, M., Mueller, R.D., Nolan, T., Pfaffl, M.W. 2013. Guidelines for Minimum Information for Publication of Quantitative Digital PCR Experiments. Clinical Chemistry 59(6): 892–902.

International Organization for Standardization (ISO). 1994a. ISO 5725-2:1994. Accuracy (Trueness and Precision) of Measurement Methods and Results – Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method. Geneva, Switzerland.

International Organization for Standardization (ISO). 1994b. ISO 5725-3:1994. Accuracy (Trueness and Precision) of Measurement Methods and Results – Part 3: Intermediate Measures of the Precision of a Standard Measurement Method. Geneva, Switzerland.

International Organization for Standardization (ISO). 1994c. ISO 5725-1:1994. Accuracy (Trueness and Precision) of Measurement Methods and Results – Part 1: General Principles and Definitions. Geneva, Switzerland.

International Organization for Standardization (ISO). 2000. ISO/TR 13843:2000(E). Water Quality – Guidance on Validation of Microbiological Methods. Geneva, Switzerland.

Ishii, S., Ksoll, W.B., Hansen D.L., Hicks, R.E., Sadowsky, M.J. 2006. Poster: Source Tracking of *Escherichia coli* at the Duluth Boat Club Beach: Nearshore Sand Acts as a Temporal Source and Sink of this Fecal Indicator Bacterium. American Society for Microbiology 106th General Meeting. ASM. Orlando, FL.

Isobe, K.O., Tarao, M., Chiem, N.H., Minh, L.Y., Takada, H. 2004. Effect of Environmental Factors on the Relationship between Concentrations of Coprostanol and Fecal Indicator Bacteria in Tropical (Mekong Delta) and Temperate (Tokyo) Freshwaters. Applied and Environmental Microbiology 70(2): 814–821.

Jenkins, M.B., Endale, D.M., Fisher, D.S., Adams, M.P., Lowrance, R., Newton, G.L., Vellidis, G. 2012. Survival Dynamics of Fecal Bacteria in Ponds in Agricultural Watersheds of the Piedmont and Coastal Plain of Georgia. Water Research 46: 176–186.

Khan, I.U., Gannon, V., Kent, R., Koning, W., Lapen, D.R., Miller, J., Neumann, N., Rhillips, R., Robertson, W., Topp, E., van Bochove, E., Edge, T.A. 2007. Development of a Rapid Quantitative PCR Assay for Direct Detection and Quantification of Culturable and Non-culturable *Escherichia coli* from Agriculture Watersheds. Journal of Microbiological Methods 69(3): 480–488.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., Wuertz, S. 2007. 16S rRNA-based Assays for Quantitative Detection of Universal, Human-, Cow-, and Dog-specific Fecal Bacteroidales: A Bayesian Approach. Water Research 41(16): 3701–3715.

Lavender, J.S. and Kinzelman, J.L. 2009. A Cross Comparison of QPCR to Agar-based or Defined Substrate Test Methods for the Determination of *Escherichia coli* and Enterococci in Municipal Water Quality Monitoring Programs. Water Research 43(19): 4967–4979.

Lee, J.L. and Deininger, R.A. 2004. Detection of *E. coli* in Beach Water within 1 Hour Using Immunomagnetic Separation and ATP Bioluminescence. Luminescence 19: 31–36.

Lee, C.M., Griffith, J.F., Kaiser, W., Jay, J.A. 2010. Covalently Linked Immunomagnetic Separation/Adenosine Triphosphate Technique (Cov-IMS/ATP) Enables Rapid, in-field Detection and Quantification of *Escherichia coli* and *Enterococcus* spp. in Freshwater and Marine Environments. Journal of Applied Microbiology 109(1):324–3.

Liang, Z., He, Z., Zhou, X., Powell, C.A., Yang, Y., Roberts M.G., Stoffella, P.J. 2012. High Diversity and Differential Persistence of Fecal Bacteroidales Population Spiked into Freshwater Microcosm. Water Research 46: 247–257.

Litton, R.M., Ahn, J.H., Sercu, B., Holden, P.A., Sedlak, D.L., Grant, S.B. 2010. Evaluation of Chemical, Molecular, and Traditional Markers of Fecal Contamination in an Effluent Dominated Urban Stream. Environmental Science and Technology 44(19): 7369–7375.

Lund, M., Nordentoft, S., Pedersen, K., Madsen, M. 2004. Detection of *Campylobacter* spp. in Chicken Fecal Samples by Real-Time PCR. Journal of Clinical Microbiology 42: 5125–5132.

McCarthy, D.T., Hathaway, J.M., Hunt, W.F., Deletic, A. 2012. Intra-event Variability of *Escherichia coli* and Total Suspended Solids in Urban Stormwater Runoff. Water Research 46(20): 6661–6670.

Noble, R.T., Blackwood, A.D., Griffith, J.F., McGee, C.D., Weisberg, S.B. 2010. Comparison of Rapid Quantitative PCR-Based and Conventional Culture-based Methods for Enumeration of *Enterococcus* spp. and *Escherichia coli* in Recreational Waters. Applied and Environmental Microbiology 76(22): 7437–7443.

Paar, J., Varma,M., Siefring, S., Oshima, K., Haugland, R.A. (2014) Development and evaluation of a culture-independent method for source determination of fecal wastes in surface and storm waters using reverse transcriptase-PCR detection of FRNA coliphage genogroup gene sequences. (submitted)

Rogers, S.W., Donnelly, M., Peed, L., Kelty, C.A., Mondal, S., Zhong, Z., Shanks, O.C. 2011. Decay of Bacterial Pathogens, Fecal Indicators, and Real-time Quantitative PCR Genetic Markers in Manure-Amended Soils. Applied and Environmental Microbiology 77(14): 4839–4848.

Russell, T.L., Sassoubre, L.M., Wang, D., Masuda, S., Chen, H., Soetjipto, C., Hassaballah, A., Boehm, A.B. 2013. A Coupled Modeling and Molecular Biology Approach to Microbial Source Tracking at Cowell Beach, Santa Cruz, CA, USA. Environmental Science & Technology 47: 10231-10239.

Sauer, E.P., Vandewalle, J.L., Bootsma, M.J., McLellan, S.L. 2011. Detection of the Human Specific *Bacteroides* Genetic Marker Provides Evidence of Widespread Sewage Contamination of Stormwater in the Urban Environment. Water Research 45: 4081–4091.

Schoen, M.E., Soller, J.A., Ashbolt, N.J. 2011. Evaluating the Importance of Faecal Sources in Human-impacted Waters. Water Research 45: 2670–2680.

Sidhu, J.P., Hodgers, L., Ahmed, W., Chong, M.N., Toze, S. 2012. Prevalence of Human Pathogens and Indicators in Stormwater Runoff in Brisbane, Australia. Water Research 46(20): 6652–6660.

Solo-Gabriele, H.M., Wolfert, M.A., Desmarais, T.R., Palmer, C.J. 2000. Sources of *Escherichia coli* in a Coastal Subtropical Environment. Applied and Environmental Microbiology 66(1): 230–237.

Sonu, C.J. 2012. Field Observation of Nearshore Circulation and Meandering Currents. Journal of Geophysical Research 77(18): 3232–3247.

Stewart, J.R., Gast, R. J., Fujioka, R.S., Solo-Gabriele, H.M., Meshke, J.S., Amaral-Zettler, L.A., del Castillo, E., Polz, M.F., Collier, T. K., Strom, M.S., Sinigaliiano, C.D., Moeller, P.D.R., Holland, A.F. 2008. The Coastal Environment and Human Health: Microbial Indicators, Pathogens, and Reservoirs. Environmental Health 7(Suppl2): S3.

Stumpf, C.H., Piehler, M.F., Thompson, S., Noble, R.T. 2010. Loading of Fecal Indicator Bacteria in North Carolina Tidal Creek Headwaters: Hydrographic Patterns and Terrestrial Runoff Relationships. Water Research 44: 4704–4715.

Taniguchi, M. 2002. Tidal effects on submarine groundwater discharge into the ocean. Geophysical Research Letters 29(12): 1–3.

Telech, J.W., Brenner, K.P., Haugland, R., Sams, E., Dufour, A.P., Wymer, L., Wade, T.J. 2009. Modeling *Enterococcus* Densities Measured by Quantitative Polymerase Chain Reaction and Membrane Filtration Using Environmental Conditions at Four Great Lakes Beaches. Water Research 43(19): 4947–4955.

U.S. EPA. 1986. EPA's Ambient Water Quality Criteria for Bacteria – 1986. U.S. Environmental Protection Agency: Washington, D.C. EPA440/5-84-002.

U.S. EPA. 2001a. Method 1601: Male-specific (F+) and Somatic Coliphage in Water by Two-step Enrichment Procedure. U.S. Environmental Protection Agency: Washington, D.C. EPA 821-R-01-030. Available at: http://www.epa.gov/nerlcwww/documents/1601ap01.pdf

U.S. EPA, 2001b. Method 1602: Male-specific (F+) and Somatic Coliphage in Water by Single Agar Layer (SAL) Procedure. U.S. Environmental Protection Agency: Washington, D.C. EPA 821-R-01-029. Available at: http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/2008_11_25_methods_method_biological_1602.pdf

U.S. EPA. 2002. Method 1604: Total Coliforms and *Escherichia coli* in Water by Membrane Filtration Using a Simultaneous Detection Technique (MI Medium). Office of Water EPA 821-R-02-024. Available at: http://www.epa.gov/microbes/documents/1604sp02.pdf

U.S. EPA. 2004. Water Quality Standards for Coastal and Great Lakes Recreation Waters, Final Rule. Federal Register, 69(220): 67218–67243.

U.S. EPA. 2005. Manual for the Certification of Laboratories Analyzing Drinking Water. EPA 815R-05-004.

U.S. EPA. 2009a. Method 1600: Enterococci in Water by Membrane Filtration Using Membrane-*Enterococcus* Indoxyl-β-D-Glucoside Agar (mEI). U.S. Environmental Protection Agency: Washington, D.C. EPA 821-R-02-022. Available at: http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/method_1600.pdf

U.S. EPA. 2009b. Method 1603: *Escherichia coli* (*E. coli*) in Water by Membrane Filtration Using Modified membrane-Thermotolerant *Escherichia coli* Agar (Modified mTEC). U.S. Environmental Protection Agency: Washington, D.C. EPA-821-R-09-007. Available at: http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/method_1603.pdf

U.S. EPA. 2009c. Method Validation of U.S. Environmental Protection Agency Microbiological Methods of Analysis. Forum on Environmental Measurements (FEM) Document Number 2009-01. Available at: http://www.epa.gov/fem/pdfs/final_microbiology_method_guidance_110409.pdf

U.S. EPA. 2010a. EPA Microbiological Alternate Test Procedure (ATP) Protocol for Drinking Water, Ambient Water, Wastewater, and Sewage Sludge Monitoring Methods. EPA-821-B-10-001.

U.S. EPA. 2010b. Report on 2009 National Epidemiologic and Environmental Assessment of Recreational Water Epidemiology Studies (NEEAR 2010 – Surfside & Boquerón). EPA-600-R-10-168. Available at: http://www.epa.gov/neear/files/Report2009v5_508comp.pdf

U.S. EPA. 2012a. Recreational Water Quality Criteria. Office of Water 820-F-12-058.

U.S. EPA. 2012b. Method 1611: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay. Office of Water EPA-821-R-12-008. Available at: http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/Method-1611-Enterococci-in-Water-by-TaqMan-Quantitative-Polymerase-Chain-Reaction-qPCR-Assay.pdf

U.S. EPA 2014. Overview of Technical Support Materials: A Guide to the Site-Specific Alternative Criteria TSM documents. EPA-820-R-14-010.

U.S. FDA, Center for Drug Evaluation and Research (CDER). 2001. Guidance for Industry. Bioanalytical Method Validation. Available at: http://www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf

Viau, E.J., Goodwin, K.D., Yamahara, K.M., Layton, B.A., Sassoubre, L.M., Burns, S.L., Tong, H.I., Wong, S.H.C., Lu, Y., Boehm, A.B. 2011. Bacterial Pathogens in Hawaiian Coastal Streams—Associations with Fecal Indicators, Land Cover, and Water Quality. Water Research 45: 3279–3290.

Wade, T.J., Calderon, R.L., Brenner, K.P., Sams, E., Beach, M., Haugland, R., Wymer, L., Dufour, A.P. 2008. High Sensitivity of Children to Swimming-associated Gastrointestinal Illness – Results Using a Rapid Assay of Recreational Water Quality. Epidemiology 19(3): 375–383.

Wade, T.J., Sams, E., Brenner, K.P., Haugland, R., Chern, E., Beach, M., Wymer, L., Rankin, C.C., Love, D., Li, Q., Noble, R., Dufour, A.P. 2010. Rapidly Measured Indicators or Recreational Water Quality and Swimming-associated Illness at Marine Beaches: A Prospective Cohort Study. Environmental Health 9: 66.

Walters, S.P., Yamahara, K.M., Boehm, A.B. 2009. Persistence of Nucleic Acid Markers of Health-relevant Organisms in Seawater Microcosms: Implications for their Use in Assessing Risk in Recreational Waters. Water Research 43: 4929–4939.

Whitman, R.L., and Nevers, M.B. 2003. Foreshore Sand as a Source of *Escherichia coli* in Nearshore Water of a Lake Michigan Beach. Applied and Environmental Microbiology 69(9): 5555–5562.

Whitman, R.L., Shively, D.A., Pawlik, H., Nevers, M.B., Byappanahalli, M.N. 2003. Occurrence of *Escherichia coli* and Enterococci in *Cladophora* (Chlorophyta) in Nearshore Water and Beach Sand of Lake Michigan. Applied and Environmental Microbiology 69(8): 4714–4719.

Whitman, R.L., Ge, Z., Nevers, M.B., Boehm, A.B., Chern, E.C., Haugland, R.A., Lukasik, A.M., Molina, M., Przybyla-Kelly, K., Shively, D.A., White, E.M., Zepp, R.G., Byappanahalli, M.N. 2010. Relationship and Variation of qPCR and Culturable Enterococci Estimates in Ambient Surface Waters Are Predictable. Environmental Science and Technology 44(13): 5049–5054.

Willmott C.J., and Wicks, D.E. 1980. An Empirical Method for the Spatial Interpolation of Monthly Precipitation within California. Physical Geography 1(1): 59–73.

Willmott C.J., Robeson, S.M., Matsuura, K. 2011. A Refined Index of Model Performance. International Journal of Climatology 32(13): 2088–2094.

World Health Organization (WHO). 2003. Assessing Microbial Safety of Drinking Water: Improving Approaches and Methods. Published on behalf of WHO and the Organisation for Economic Cooperation and Development (OECD) by IWA Publishing. Available at: http://www.who.int/water_sanitation_health/dwq/9241546301full.pdf

Wymer, L.J., Brenner, K.P., Martinson, J.W., Stutts, W.R., Schaub, S.A., Dufour, A.P. 2005. The EMPACT beaches project, Results from a Study on Microbiological Monitoring in Recreational Waters. EPA 600/R-04-023. http://water.epa.gov/type/oceb/beaches/upload/2006_06_21_microbes_empact-3.pdf

Wymer, L.J. and Wade, T.J. 2007. The Lognormal Distribution and Use of the Geometric Mean and the Arithmetic Mean in Recreational Water Quality Measurement. In Statistical Framework for Recreational Water Quality Criteria and Monitoring. New Jersey, John Wiley & Sons, Inc.

Yamahara, K.M., Layton, B.A., Santoro, A.E., Boehm, A.B. 2007. Beach Sands along the California Coast are Diffuse Sources of Fecal Bacteria to Coastal Waters. Environmental Science & Technology 41: 4515–4521.

Yamahara, K.M., Walters, S.P., Boehm, A.B. 2009. Growth of Enterococci in Unaltered, Unseeded Beach Sands Subjected to Tidal Wetting. Applied and Environmental Microbiology 75(6): 1517–1524.

Yamahara, K.M., Sassoubre, L.M., Goodwin, K.D., Boehm, A.B. 2012. Occurrence and Persistence of Bacterial Pathogens and Indicator Organisms in Beach Sand along the California Coast. Applied and Environmental Microbiology 78(6): 1733–1745.

Appendix A: Factors that Determine Whether to Pursue Comparison of Indicator/Methods

Appendix A summarizes background information to help you anticipate whether a correlation between indicator organisms is reasonable to expect at a particular site and what site features you should consider when designing sampling plans. After considering the information in this appendix and how it applies to the site, you should be able to decide if you should proceed to the steps outlined in Section 2. Because each site is unique, you should consider the weight-of-evidence regarding these factors for your site. The relationships between the densities of pairs of fecal indicator organisms at a site are likely to be highly site specific, and you can establish them only by using site-specific data. In addition, after you have conducted water quality monitoring, you can use this information to help you understand the results of the comparison of two indicator/methods.

Factors that influence the correlation between indicator/method pairs are described in the following four subsections, type of assay, fecal sources, age and proximity of fecal sources, and hydrometeorological[22] factors.

Type of Assay

One of the most recognized differences between methods is the difference between methods that require microorganisms to grow (culture methods) and methods that detect DNA (molecular methods) (Converse et al., 2012). Even if both methods detect the same microorganism, because of the technological specificities of the two methods, they really do not measure the same thing. Predicting that qPCR methods (molecular methods) will track with culture-based methods requires understanding what each type of method includes and excludes. At a simplistic level, culture methods detect any microbe that can grow in the medium used for the assay. Historically, scientists classified microorganisms by the media supporting their growth. For example, molecular methods detect DNA or ribonucleic acid (RNA) sequences and, depending on the specific design of the method, can detect a more limited or a broader group of microorganisms than a given culture method.

Khan et al. (2007) compared qPCR enumeration of *E. coli* in waters of agriculture-dominated watersheds against enumeration with membrane filtration (culture) methods. They, like Haugland et al. (2005), found that qPCR-based enumerations yielded consistently higher estimates of density than culture methods (attributed to lack of discrimination between DNA from live and dead cells) and that standard curves (in this case, based on both dilution water and autoclaved agricultural water) had high coefficients of variation. Converse et al. (2009) showed a linear relationship between the density of culture-enumerated enterococci and the density of qPCR-enumerated enterococci. In that study, results of the methods differ by several orders of magnitude at low log-densities and are of comparable magnitude at high densities.

[22] Hydrometeorology includes the disciplines of both meteorology and hydrology. It is the study of the transfer of water and energy between the land surface and the lower atmosphere.

Whitman et al. (2010) found a general trend toward a linear relationship between log-transformed qPCR and culture counts of enterococci for data collected at 37 marine and fresh water sites in the United States. Site factors that seemed to favor a consistent relationship between culture and qPCR counts were the presence of a point source of fecal contamination and relatively high indicator organism levels. The latter site factor might relate to the high variability in qPCR counts at low indicator organism levels. Based on trends in the covariation in qPCR and culturable enterococci over all sites studied, the authors suggest that a linear relationship adequately describes the covariation in log-transformed indicator organism counts of *Enterococcus* via culture and qPCR methods.

Converse et al. (2012) compared the agreement between EPA method 1600 and qPCR-measured enterococci between and among three beaches. They found that correlations were stronger with samples collected in the mornings compared to afternoons and at samples collected at beaches with more concentrated sources compared to diffuse sources. They also found that the ratio of the two measures varied between beaches.

Throughout the following sections, we note differences and similarities between culture and molecular methods.

Predominant fecal pollution source

In addition to being present in human fecal material, FIB are associated with a variety of non-human sources, such as animal waste and non-fecal environmental sources (Stewart et al., 2008; Fujioka and Byappanahalli, 2003; Byappanahalli et al., 2011). For example, enterococci and *E. coli* can be indigenous, autochthonous members of the microbial community in waterbodies, sands, sediments, soils, or plants. Thus FIB can come from different sources at different sites. You should understand which fecal and environmental sources contribute FIB to your site. If the two indicator/methods you are comparing are likely to detect FIB from different sources, the two indicator/methods are less likely to yield results that correlate over time, because of the different composition of microflora from differing sources.

The mixture of indicator organisms at a particular site is variable; it is a result of the net loading of the indicator organism from all fecal pollution sources, loading of the indicator organism from non-fecal sources, and growth or decline (via die-off, predation, sedimentation, and other processes) at the site. Figure A-1 depicts one possible scenario. If the predominant fecal pollution source has a characteristic ratio of the abundance of Indicator 1 to that of Indicator 2, the density of Indicator 2 should generally increase as the density of Indicator 1 increases.[23] At many sites, the relative loading of the various fecal and non-fecal sources can change over time. If the different sources have different characteristic ratios of the indicators, the result would be either different relationships between indicator densities over time, or an apparent poor correlation between the indicators when all data are considered. Also, in the scenario in Figure A-1, if the predominant source is intermittent (e.g., storm flow), other sources might

[23] In Figure A-1, Indicator 1 and Indicator 2 refer to an indicator and its associated enumeration method. In some cases, you could enumerate the same organism by different methods. For example, Indicator 1 could be enterococci by culture method, and Indicator 2 could be enterococci by qPCR.

dominate when the primary source recedes. Alternatively, the predominant source of fecal pollution might contain only one of the two indicators you are comparing. In this case, indicators that are from different sources would be unlikely to correlate. In addition, the scenario depicted in Figure A-1 is for two different fecal indicator organisms, but you can compare two different methods that measure the same organism with the approach in this TSM.

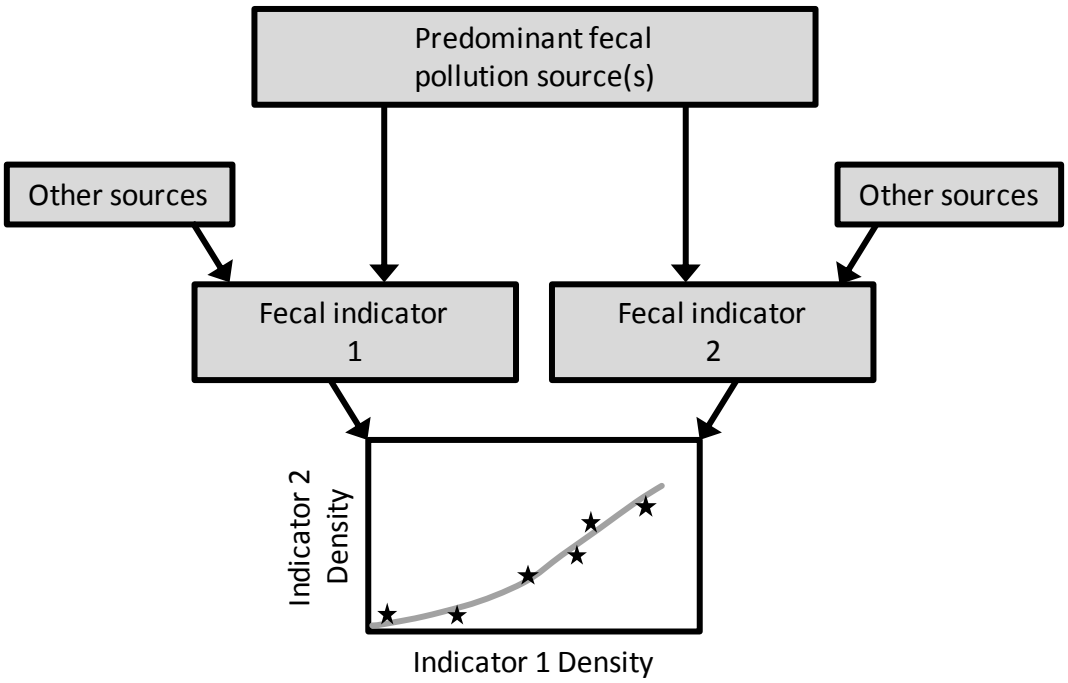

Figure A-1. Idealized relationship between two fecal indicators as the result of their net loading from all sources

Schoen et al. (2011) evaluated the relationship between indicator organisms and contributions from fecal pollution sources using a quantitative microbial risk assessment (QMRA) approach. They simulated a hypothetical site with fecal indicator loading from multiple sources (treated and untreated sewage, sediments, and livestock wastes). In the QMRA scenarios, the waterbody indicator density and the illness level were both held constant to compare which ratios of sources could yield the particular indicator density and illness level combination. Enterococci density assayed by culture was dominated by the untreated (or poorly treated) sewage. In contrast, *Enterococcus* density assayed by qPCR was dominated by the secondary-treated disinfected municipal wastewater effluent. This finding is consistent with the very high densities of viable enterococci typical of raw sewage and the simultaneous presence of very high levels of qPCR signal and very low levels of culture signal typical of disinfected waters. This finding also shows that strong correlations between different indicator/methods are unlikely if contributions from sources with different characteristics change over time. Sidhu et al. (2012) found the abundance of indicator organisms in water varied widely during and between storms, but that *E. coli* were generally less abundant than enterococci (both measured by culture

methods) and that human-specific Bacteroidales were present in many, but not all samples. The existence of a general relationship between enterococci and *E. coli* (e.g., Dickenson and Sansalone, 2011) suggests the possibility of a relationship between fecal indicator organism counts in run-off, although many factors can influence that relationship. For example, differences in the degree to which the fecal indicator organisms associate with particles and differences in microbial die-off between storms can affect the relationship.

The authors also noted they collected and concentrated relatively large (20 liter) samples prior to qPCR analyses for Bacteroidales enumeration. qPCR inhibition is possible when analyzing concentrated samples, which is an additional reason FIB counts might not correlate well, particularly at low FIB densities (Sidhu et al., 2012).

The relative abundances of FIB at sites that run-off has impacted are particularly difficult to generalize. Multiple reports (Sauer et al., 2011; McCarthy et al., 2012) have demonstrated that human-specific fecal pollution markers are frequently present in urban run-off, but the markers can be present only intermittently. The relative FIB population in run-off can change with time, resulting in poor correlation between human-specific markers and traditional FIB, as Sauer et al. (2011) observed. In that study, moderate to low levels of *E. coli* (via culture methods) frequently were associated with high levels of human *Bacteroides*, and moderate to low levels of human *Bacteroides* often were associated with high levels of *E. coli*. Lavender and Kinzelman (2009) found no significant difference in qPCR counts of enterococci for dry- and wet-weather discharges, despite a significant difference in culture counts of the enterococci; they also found no significant correlation (on the basis of R-squared) between the qPCR and culture counts of enterococci in wet-weather run-off. The authors speculated that wet weather influenced culture counts more than qPCR because the fresh, viable FIB loading from wet weather can be large compared with the dry (background) level. In contrast, the background level of indicator organism DNA from viable, nonviable, and extracellular DNA was relatively high and the load of new indicator organism DNA associated with rain events was small relative to background levels (Lavender and Kinzelman, 2009).

The specificity of alternative indicators/methods can strongly influence their correlation at a particular site. Converse et al. (2009) conducted laboratory experiments examining the abundance of *Bacteroides* spp. (measured via qPCR) and enterococci (measured by qPCR and culture methods) in samples spiked with human sewage and gull guano. They found that enterococci were plentiful in samples spiked with both human sewage and gull guano, whereas *Bacteroides* were plentiful only in the sewage-spiked samples. This finding indicates that the site-specific fecal pollution source could have a strong influence on the relationship between a traditional and alternative indicator/method at a site, particularly when the indicator/methods have differing specificities. Changes in the contributions from different fecal pollution sources, if possible at the site, can further influence the relationship between the indicator/methods. Another study (Litton et al., 2010) on the relative abundance (in terms of loads) of specific markers and traditional FIB similarly found that the traditional FIB are ubiquitous and only tenuously related to predominant fecal pollution sources, whereas specific markers occur less generally and their loading differs substantially from that of traditional FIB. These observations

imply that a strong correlation might not exist between the specific indicator/method and the traditional FIB at a given site.

Age and proximity of fecal pollution sources

At a given site or in a given source, environmental survival can vary widely among fecal indicator organisms. This variation influences the comparability between indicator/methods at a given downstream site because of differential fate and transport processes, including attenuation, predation, injury, sedimentation, and attachment. You should know enough about the fecal source dynamics at the site to determine how important fate and transport of fecal material is at the site.

Many studies compare molecular methods such as qPCR to culture methods. Important to note is that the qPCR signal usually is a measure of the abundance of viable cells, nonviable cells, and extracellular DNA. Therefore, one reason that rates in qPCR signal decay might differ from those of culturable organisms is that culture methods discriminate between viable and viable but non-culturable cells (i.e., injured cells), and qPCR does not.

Microcosm experiments such as those Walters et al. (2009) conducted have demonstrated widely different environmental decay rates for different indicator organisms and among targets of qPCR and culture assays. Liang et al. (2012) conducted microcosm studies in fresh water seeded with bovine or human feces and determined decay rates of *E. coli* (as determined via culture methods), the cattle Bacteroidales DNA marker CF183, and the human Bacteroidales DNA and RNA markers for HF183. The first-order decay rates for the CF183 and HF183 markers were similar at around 0.7 natural logs per day, whereas the culture count of *E. coli* declined at a rate of around 0.3 per day for the microcosms seeded with human and bovine feces. Yamahara et al. (2012) observed the decay of enterococci (via culture and qPCR), F+ phages, Bacteroidales (via qPCR), and other indicator organisms and pathogens in microcosms of unaltered beach sand seeded with sewage. Yamahara et al. (2012) found similar first-order decay rates for FIB and pathogens, but widely different decay rates for qPCR and culture targets for the same organism. For example, the culture count of enterococci decayed nearly three times faster than the qPCR target level for enterococci.

Bae and Wuertz (2009) observed much faster decay among Bacteroidales when using qPCR with propidium monoazide (PMA)[24] than when using qPCR without PMA. The host-specific 2-log reduction time in microcosm experiments differed by more than a factor of 5 for cells, as determined by qPCR with PMA, and DNA, as determined by qPCR without PMA. These findings relate directly to the relationship between counts of different indicator/methods at a particular site. When comparing conventional qPCR results with culture data, the correlation will reflect the fact that the assays are counting two different targets whose environmental decay rates might be very different.

Expecting growth to differ among fecal organisms and to cause additional differences in the relationship between levels of indicator organisms is reasonable. Yamahara et al. (2012)

[24] PMA is a compound that binds to DNA from dead cells and extracellular DNA and prevents replication.

observed possible growth of enterococci (as measured by culture methods) in microcosms of unaltered beach sands seeded with sewage. Yamahara et al. (2009) also observed evidence of growth in microcosm experiments of intermittent wetting of unaltered beach sands for both qPCR and culture levels of enterococci; a time series of culture and qPCR counts appeared to indicate greater change (growth) in the target of the culture method than in the qPCR target. Researchers have documented fecal indicator organism growth in ponds and flowing waters (Carrillo et al., 1985; Davies et al., 1995; Isobe et al., 2004; Jenkins et al., 2012), soils, sands and sediments (Hardina and Fujioka, 1991; Whitman and Nevers, 2003; Alm et al., 2006; Byappanahalli et al., 2006; Ishii et al., 2006), algae (Whitman et al., 2003), and on submerged aquatic vegetation (Badgley et al., 2010).

The same basic factors—temperature, sunlight, predation, salinity, and moisture (for organisms in soils and sands)—appear to govern the survival of fecal indicator organisms. These factors, however, have different influences among different indicator organisms and at different sites. The literature widely reports the influence of these factors on survival of the traditional, culture-based fecal indicator/methods. Much less data are available for the decay in qPCR counts and for less well-studied indicators such as *Bacteroides*, although several recent studies have produced pertinent data. For example, Bell et al. (2009) conducted experiments in microcosms of unfiltered and filtered stream waters to determine decay/removal rates for *Bacteroides* 16S ribosomal RNA (rRNA) genes derived from equine fecal samples. The authors found that predation plays a significant role in decline of *Bacteroides* over time and that temperature is the primary independent variable governing the decay of *Bacteroides* in the waters they studied.

Survival among indicator organisms differs along the exposure pathway, and in the water column, sands, and sediments at a particular site. A pair of indicator organisms with a characteristic relative abundance in fresh fecal deposits would have a different relative abundance in run-off if their survival rate in manures differs from that in water. Rogers et al. (2011) studied the rate of decline in FIB, host-specific genetic (qPCR) markers, and bacterial human pathogens in soils amended with cattle and swine manure. They found relatively poor correlation between qPCR and culture (MPN) counts of enterococci and *E. coli*. They also found that host-associated qPCR genetic markers for microbial source tracking decayed rapidly to non-detectable levels long before FIB did. The authors further concluded that, even though host-associated qPCR genetic markers are good indicator/methods of point source or recent nonpoint source fecal contamination, they might not be reliable for nonpoint source fecal contamination events that occur weeks after manure application on land.

Although the basis of this assessment is comparison of genetic marker survival with pathogen survival, the ratio of genetic marker abundance to FIB abundance will differ markedly for fresh and aged manures. Thus, fecal pollution age could cause greater variability in the relative abundance of different indicator/methods and negatively impact the correlation between the indicator/methods at a receiving site.

Meteorological and hydrological features

Hydrometeorological site characteristics influence the fecal source dynamics of waterbodies. You should know the hydrometeorological patterns of the site so you can design a sampling plan that captures the typical variability for the site. The hydrometeorological factors may also help you understand why two indicator/methods do not correlate.

Numerous studies have shown that run-off impacts recreation sites sporadically and can be associated with drastic water quality changes during a rain event or between rain events (e.g., Stumpf et al., 2010; McCarthy et al., 2012; Sidhu et al., 2012). McCarthy et al. (2012) found that catchment size, drainage infrastructure complexity, and land use influenced total suspended solids and *E. coli* density, variability, rates of density change, and strength of the first flush (the fraction of the full load appearing during the first flush). Catchment characteristics appear to influence the variability of *E. coli* (and presumably other indicator organisms). Thus, drainage to the site influences the relationship between indicator/methods when run-off is a significant source. These differences in sites further highlight why this TSM focuses on site-specific relationships between indicator/methods.

Byappanahalli et al. (2010) assessed how site hydrometeorological factors influence the occurrence of culture and qPCR measures of the same fecal organism. The hydrometeorological factors predicting the density of *Enterococcus* CFU were similar to those predicting *Enterococcus* cell equivalents (CE) by qPCR. For culture counts of *Enterococcus* at a beach, discharge of a nearby stream, wind direction, and lake turbidity were the best predictors. For qPCR counts of *Enterococcus*, the best predictors were discharge from the nearby stream and the product of turbidity and wave height. The predictive factors explained more of the variability in the culture counts of *Enterococcus* than the qPCR counts, perhaps due in part to the high variability associated with low qPCR counts of *Enterococcus*. Interestingly, despite their being predicted by similar hydrometeorological factors, the culture and qPCR counts of *Enterococcus* correlated poorly, and the ratio of the mean CFU count to the mean qPCR count varied widely among sample locations.

A similar study (Telech et al., 2009) used multiple linear regression with backward elimination to identify the site, sample, and meteorological features that explain the variance in observed FIB densities (as measured by culture and qPCR techniques) at four beaches. The explanatory factors sometimes differed for the same indicator organism as measured by the two methods. In some cases, the sign of the correlation between the factor and the indicator organism density differed for the two analytical targets. For example, at one beach, the correlation between wave height and *Enterococcus* density as measured by qPCR was positive, but was negative with culture *Enterococcus* density. No single site feature consistently predicted both culture and qPCR density for all sites. The coefficients of determination (R-squared) for the final models for predicting culture and qPCR densities varied widely (range 0.22–0.94) among the beaches. Coefficients of determination were lower for the qPCR models than for the culture models for all beaches.

Lavender and Kinzelman (2009) determined the hydrometeorological factors that influenced correlation in qPCR and culture counts of *E. coli* and *Enterococcus*. Based on those factors, the authors developed and applied corrections to (subtracted from) qPCR counts to reconcile those counts with the culture counts. This approach led to very good correlation between CFUs and corrected qPCR counts (CE) for riverine and coastal (beach) sites. Although the hydrometeorological factors that necessitated a correction differed among sites, the occurrence of rainfall within 48 hours before sample collection and wave height at the time of sample collection were the factors most frequently encountered. We note that, as described above, rainfall and wave action can change the proportion of indicator organisms loaded to sites from different fecal pollution sources. For example, after rainfall when storm drains are discharging, the relative abundances of culture and qPCR indicator/methods would be closer to the abundances typical of stormwater, whereas during dry conditions the ratio could be more typical of other sources.

The effect of tides of fecal indicator bacterial densities has been documented at a number of marine beaches (i.e., Boehm and Weisberg 2005, Solo-Gabriele et al. 2000). Boehm et al. (2002) showed that bacterial densities were significantly higher during spring compared to neap tides at Huntington Beach, California. Boehm and Weisberg (2005) extended the analysis to consider the effect of tides at 60 different beaches in southern California. They found that enterococci densities were significantly higher during spring compared to neap tides at 50 of the 60 beaches. The highest enterococci densities were observed during spring-ebb tides. The cause of the tidal effects appears to depend on the specific beaches. Various processes at marine beaches are modulated by the tide including (1) submarine groundwater discharge (Taniguchi, 2002), (2) washing of contaminated sand (Yamahara et al., 2007) or wrack (Russell et al., 2013) on the beach face , (3) discharge from lagoons and marshes (Grant et al., 2001), and (4) nearshore currents (Sonu, 1972).

In conclusion, multiple site-specific and assay-specific factors contribute to whether indicator/methods will correlate. These include: the assay-specific factors for methods you are comparing, fecal sources contributing to FIB at the site, the proximity of fecal sources to the site, and the hydrometeorological influences at the site. Investigation of these factors can help you decide if a correlation might be likely or unlikely before water quality sampling is undertaken. In addition, these factors might improve your understanding of the results of your water quality study.

Appendix B: Example Sampling and Analysis Plan

We adapted this example SAP from the City of Racine Health Department's Sampling Manual (Racine, Wisconsin). You are not limited to the elements shown in this example; you may use other SAP formats and designs. You should tailor your SAP to the location you are sampling. Example procedures for beaches and rivers are described below. Mention of specific products is not an endorsement. You can use whatever products you typically would use for sampling. This example SAP includes references to other City of Racine documents that are not included.

Procedure – Sample Collection for Beaches

Field Sampling Equipment & Materials

1. Equipment

a. Thermometer – Ready to use as supplied. All thermometers are calibrated against a calibrated thermometer. Refer to the *Thermometer Quality Control* procedure to calibrate thermometers.

b. Insulated Cooler – Ready to use as supplied. Do not use insulated coolers with excessive cracks, tears, or holes. Storage should be available for retaining additional supplies such as extra sample bags and writing utensils. Sanitize coolers after each use.

c. Clip Board – Ready to use as supplied.

d. Time Piece – Ready to use as supplied.

e. Waders or Wet Suit – Ready to use as supplied. The waders or wet suit should only be used when temperatures are very cold and/or the site is hazardous due to the substrate or debris. Do not use waders or the wet suit on extremely hot days due to the potential for heat stroke/exhaustion.

2. Materials

a. Whirl-Pak® Bags – Ready to use as supplied; 18oz bags. All boxes are tested for sterility. Refer to the *Container Sterility Test* procedure to testing new boxes.

b. Beach Sampling Forms – Usually customized by the State or local agency sponsoring the sampling.

c. Ice Packs – Ready to use as supplied. Foam or gel refrigerant blocks produce a more uniform temperature than blue ice. Just keep samples between 0 and 4°C. Use enough to cover the sides and bottom of the cooler.

d. Pen – Ready to use as supplied. Do not use gel pens and keep colors to blue or black ink. You should also have at least two pens, one for use and the other ask backup. Do not use pencil.

e. Permanent Marker – Ready to use as supplied. You should keep one in the bag for labeling in the field.

Procedure – Collection and Survey Preparation

1. Using a permanent marker, label the Whirl-Pak bags according to the Beach Sampling forms.
2. Record the date and collector's name on the Beach Sampling forms.
3. Check your email for the precipitation record from the Festival Hall rain gauge. Record the value on the Beach Sampling forms for Racine beaches.
4. Log on to the Internet and go to the following website to obtain the wind direction and speed: http://weather.noaa.gov/weather/current/KRAC.html. Record the wind direction (bearing if available) and speed on the Beach Sampling forms for Racine. Before exiting, record the prior 24-hr precipitation plus wind direction and speed on the sample form for Quarry.

Procedure – Sample & Site Data Collection

1. Walk to your sampling location and take the air temperature by holding the thermometer in the shade made by your body; acclimation should only take 2–3 minutes. Record on the sample forms. Note: If the probe is directly in the sunlight you will not obtain an accurate reading. Note: A single air temperature, taken at the mid-point, can be used for a contiguous stretch of shoreline (e.g., North and Zoo Beaches).
2. Place your equipment behind the berm-crest and remove the appropriately labeled bag. At this point, complete the sanitary survey for the site and record the following:
   a. Number of Gulls, number of Dogs, number of Geese, number of Bathers – in and out of water, algal presence, wave height and intensity, litter and debris, odors, dead fish, and other fields. Please attempt to complete all fields or indicate as Not Applicable or N/A.
3. After completing the survey, wade into the water to a depth of 24–30 inches; approximately waist deep. Take the water temperature by lowering the thermometer one foot below the surface of the water. Acclimation should take 1–2 minutes. Note: Do not take a temperature directly from the sample as you may subject it to contamination.
4. After taking the water temperature, prepare the Whirl-Pak bag by removing the perforated edge. Using the two white tabs, pull the mouth of the bag open; do not touch the inside of the bag. Note: Touching the mouth or interior of the bag will contaminate the sample. If this happens, return to the shore and label a new bag.
5. When the bag is open, grab the yellow tabs and turn the bag downward toward the water. In one swift motion, submerge the bag to a depth of approximately one foot while pulling up and away from your body to collect the sample. Note: Observe the direction of the waves and orient your body perpendicular to the wave direction.
6. Examine the bag to ensure enough sample has been collected; enough sample represents the bottom edge of the white labeling area. To close the bag, grab the yellow tabs and whirl the bag towards your body quickly. Twist the yellow tabs together to seal

the bag; done properly the bag will seal itself and headspace will appear. Note: If the bag is too full, use your index and thumb to pinch the approximately one inch from the top of the bag to remove sample. Note: If a bag fails to collect a sample then return to shore to label a new bag and re-collect the sample.

7. Return to the shoreline and place the sample immediately on ice; keep samples upright. Record the water temperature taken and the time the sample was collected on the sample form.
8. Examine the Beach Sample forms and fill in any missing data on the forms before leaving. Proceed to the laboratory or next sample site until all samples have been collected. Note: No more than 6 hours should elapse between sample collection and start of analysis (for bacterial culture methods).

Procedure – Sample Processing

1. Upon returning to the laboratory, place the samples in the sample refrigerator (0–4°C). Record the date and time the samples were received in the laboratory on the sample form.
2. Using either, Water Sample-MF or Water Sample – Colilert procedure, process your samples. Note: The Water Sample-MF procedure is written for preparing samples for analysis by qPCR.

Procedure – Preliminary Reporting and Data Entry

1. After processing your samples, open up the sample forms; digital copies are available on the shared drive. Complete both forms as digital copies and save as the form title plus the location and date of collection. Email the completed digital forms to the Laboratory Director. Note: Failing to complete the digital copies will delay reporting sampling results.
2. Login to the WI Beach Health website (http://www.wibeaches.us/apex/f?p=175:1). Enter the beach form data into site using the “Insert New Monitoring” and “Insert New Beach Sanitary Survey Data” options. Complete this for each site (e.g. N1-N4 and Z1-Z3). Do not enter values for *E. coli* at this time. Refer to the WI Beach Health Website Tutorial for directions and navigation.

Procedure – Sample Collection for Rivers

1. Using the map of sites and site descriptions, arrive at each location to collect a sample.
2. Upon arriving at the site take the air temperature by holding the thermometer in the shade made by your body; acclimation will occur in 2–3 minutes. Enter on report form. Note: If the probe is directly in the sunlight you will not obtain an accurate reading.
3. Remove the bag labeled with the site name, tear off the perforated edge, and pull the bag open using the white tabs. Touching only the yellow tabs, place the mouth of the bag around the holder. Wrap the snap ring around the bag and clip the ring shut so that the bag is firmly held in place; tug a couple times to make sure the bag is secure. Note:

The procedure for attaching the bag to the holder is the same for both the pole and the sampling line.

4. Remove a second unlabeled bag. This bag will be designated for taking water temperatures. Fix this bag to the sample holder by twisting the yellow tabs around the loop of the holder.
5. Extend the sampling pole or lower the sampling line to the water's surface. Submerge the bag or allow the bag to submerge itself to roughly 12 inches below the surface of the water at the fastest moving portion of the river. Fill both bags with sample and pull them back up. Note: Submerging the bag to a 12-inch depth may not be possible in all locations. If this occurs, collect as much sample as possible where water is not stagnant.
6. Using the unlabeled bag, take the water temperature; acclimation should take 1–2 minutes. Record the water temperature on the sample form.
7. Gently loosen the snap ring to release the sample being careful not to puncture the bag. Pinch the top 2 inches of the bag using your forefinger and thumb to remove some of the sample; this will provide headspace. Quickly whirl the bag toward your body and twist the ties together to seal the bag. Record the time of collection on the sample form.
8. Place the sample in the cooler in an upright position to prevent leakage. Repeat steps 1 through 8 until all samples have been collected. Note: The maximum time between sample collection and analysis should not exceed 6 hours.

Procedure – Sample Processing

1. Arriving back in the laboratory, refrigerate the samples at 0–4°C. Record the time and date the samples were received on the sample form.
2. Root River and/or Stormwater Outfall samples must have the table of tests below performed on each sample. Test should be performed in the following order due to holding time restrictions: *E. coli*, Total Chlorine, Detergents, pH, Conductivity, and Turbidity. Refer to each standard operating procedure to process samples.

Root River Samples	Stormwater Outfall Samples
Bacteria – *E. coli*, pH, Conductivity, Turbidity (NTU)	Bacteria – *E. coli*, pH, Conductivity, Turbidity (NTU), Detergents, Total Chlorine

Note: Project-specific testing or source-tracking tests may change the order in which samples must be processed. Discuss any additional testing requirements with the Laboratory Director.
NTU = nephelometric turbidity units

3. Bacteria results will not be available until the next day. All other results must be recorded on the sample form on the day of the test.

Procedure – Results Reporting and Data Entry

1. After processing all the Root River and/or Stormwater Outfall samples, open the digital version of the sample form on the shared drive. Save the form using the title, "RootRiverSamplingForm_" with the date following the underscore using the MMDDYYYY format.

2. Enter the data recorded on the sample form into the digital copy. Save the entered data and close the form until the following day when bacteria data is available. Note: If time does not allow for this step, proceed with entering all other data when bacteria results become available.
3. When bacteria results become available, record those results on the hard copy sample form and enter them into the digital copy you created the day before.
4. Email the completed digital form to the Laboratory Director and relevant contacts using the Root River Results distribution list. Place the hard copy in the River Results binder labeled for that year.

Appendix C: Explanation of Thresholds

Appendix C describes how EPA tested the approach this TSM provides by using paired ambient water quality samples from individual water bodies. EPA obtained ambient water quality data sets from EPA regional laboratories, independent researchers, the Water Environment Research Foundation, and state monitoring. Data sets were selected based on meeting all five of the following attributes: (1) paired samples with one sample being an EPA-approved method (or equivalent ATP-approved method), (2) complete raw data were provided to EPA within the 2-month period allocated for data compiling, (3) data were collected recently (after 2003), (4) there were at least 30 sample pairs above the LOQ at the site, and (5) data sets from a variety of geographical settings (east, west, coastal, inland, temperate, and tropical locations). For a list of data sets evaluated, see Tables C-1 and C-2.

The data sets were used to test various approaches for comparing paired water quality samples. Before analyses were performed, sample pairs with one or both densities below the LOQ were removed and all data were $\log_{10}$ transformed.[25] EPA chose this approach because substitution introduces a bias into the relationship between the two sets of measured values through the introduction of a dummy variable.

Section 2 of this document describes the two analyses EPA used to assess the agreement between the two indicator/methods: 1) calculate IA and 2) calculate R-squared. Tables C-1 and C-2 show the results from the example data sets.

After the data were treated to remove samples below the LOQ and $\log_{10}$ transformed, the IA was calculated. Note that other formulae can be used to calculate IA (Willmott et al., 2011, 2012), but for the purposes of this TSM, the simplest formula was adopted (Willmott and Wicks, 1980). The formula for IA is as follows:

$$IA = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2} \qquad \text{[Equation 1]}$$

where x is the EPA indicator/method and y is the alternative indicator/method, *i* is a counter, *N* is the total number of data points in the data set, $\bar{x}$ and $\bar{y}$ are the averages of the x and y data sets, respectively, and *IA* varies from 0 to 1 (Willmott and Wicks, 1980) with 1 being perfect agreement. Here x represents the EPA indicator/method (EPA method 1600, 1603, or 1611) and y the alternative indicator/method.

IA assesses both additive and proportional differences between the alternative and EPA's indicator/methods. EPA has approved other methods as "equivalent" to EPA Methods 1600 and 1603 through the ATP process. The agreement between ATP-equivalent methods and EPA-approved methods for *E. coli* and enterococci (Methods 1600 and 1603) has been previously

[25] EPA also tested non-log transformed data and treated samples below the LOQ with three different approaches: remove pairs where one or more of the indicators-methods was below the LOQ, replace samples below the LOQ with ½ LOQ, and replace samples below the LOQ with the LOQ. Results are not shown.

established and represents the best possible agreement expected between water quality measures under real conditions in actual water samples. EPA evaluated IA between EPA methods (Methods 1600 and 1603) and their corresponding ATP-equivalent assays for the various sites in the collected data set. In calculating IA for these data, the ATP-equivalent methods were considered the alternative indicator/method (y in Equation 1).

Figure C-1 shows how IA for all paired samples at the various collection sites compare. Each symbol on Figure C-1 represents one indicator/method pair at one site. Each symbol also corresponds to one row in Table C-1. The eight symbol types indicate the nature of the pairs as shown in the legend (e.g., closed squares indicate data where MF or MPN is compared to qPCR for the same organism. The open square symbols are pairs where one method is 1600 or 1603 and the other method is an ATP-equivalent method (e.g., MF compared to IDEXX).

IA values for Methods 1600 and 1603 and their ATP-equivalent counterparts collectively ranged from 0.7 to 1.0. Because IA between ATP-approved and EPA-approved methods for enterococci and *E. coli* are above 0.7, EPA considers an IA greater than or equal to 0.7 to indicate minimal additive differences. This cutoff represents the lowest IA obtained when comparing ATP-equivalent and EPA-approved indicator/method data sets from all beaches.

If IA between the alternative indicator/method and the EPA-approved indicator/method is greater than or equal to 0.7, the numerical values (both the GM and STV) for the RWQC can be directly applied to the alternative indicator/method. If the IA threshold is not passed, the R-squared value can be calculated to determine if the alternative indicator/method can be used with a new numerical limit. In this TSM, EPA is providing an option for the users to calculate R-squared on site-specific basis. The R-squared value measures proportional differences between two data sets, but not additive differences. Therefore, good agreement as determined by R-squared value suggests a new numerical limit can be calculated, but that the value will likely be different than the existing criteria. This value reflects a reasonable level of agreement, but will be evaluated on a case-by-case basis.

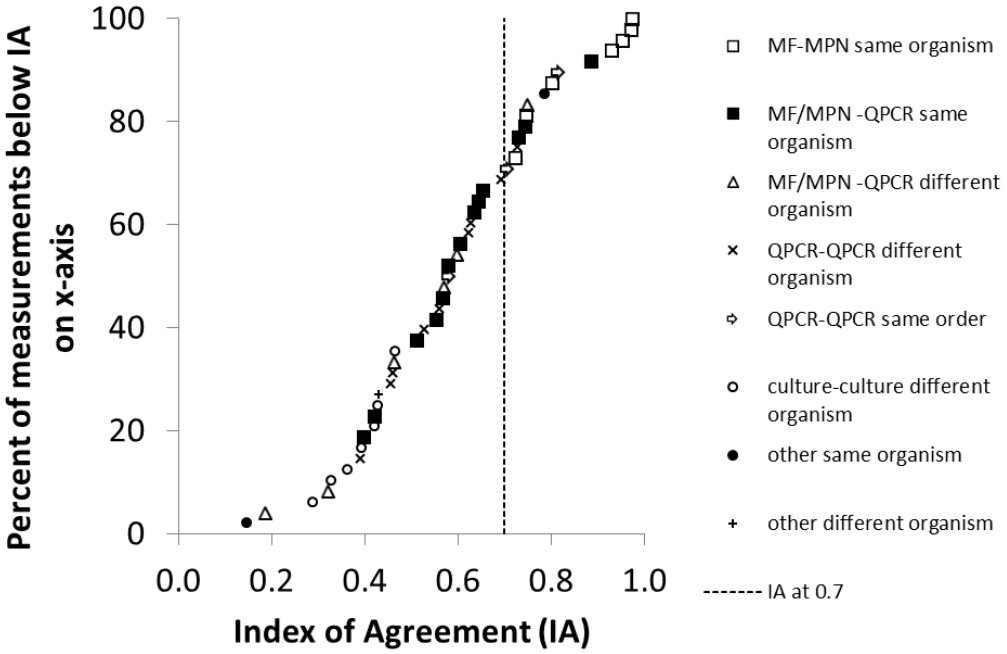

Figure C-1. Cumulative distribution of IA values coded by assay-types compared

Open squares are water quality comparisons between *E. coli* and enterococci, respectively, measured by ATP-equivalent methods and EPA-approved methods. The dotted line is at the IA threshold 0.7. Other IA values in the cumulative distribution were for comparisons of other types of water quality measures, as indicated in the legend. The 48 data sets shown each have more than 30 paired samples.

Step 2 determines if the comparison (as determined by an R-squared value) between the alternative and EPA indicator/method is fair enough that the alternative indicator/method can be used even if the agreement does not pass the IA threshold (as determined by IA). R-squared values were examined for the same ambient water data sets that were examined for IA. A threshold of >0.6 was chosen; any R-squared value obtained from a Pearson's correlation of log-transformed alternative indicator/method densities and EPA indicator/method densities must be higher than 0.6. Even though the data do not suggest a strong case for the cut-off value for R-square, EPA considers R-squared values greater than 0.6 to be indicative of agreement between two methods. EPA researchers compared methods that had good agreement and the R-squared in that case was 0.78 (Paar et al., 2014).

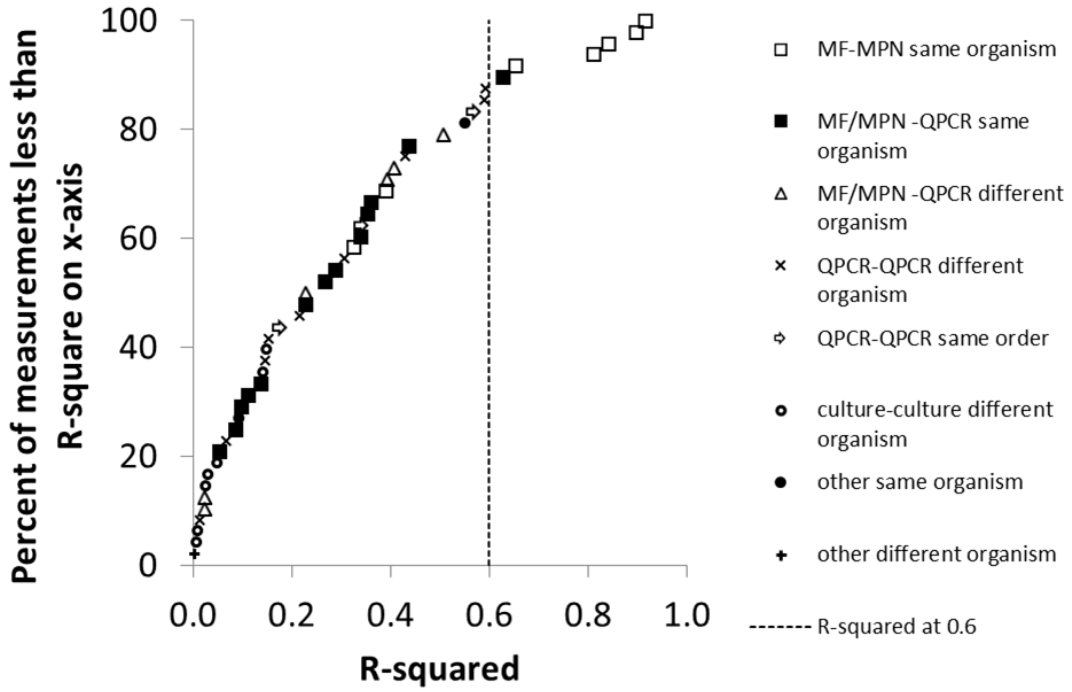

Figure C-2. Cumulative distribution of R-squared values coded by assay-types compared

Open squares are water quality comparisons between *E. coli* and enterococci, respectively, measured by ATP-equivalent methods and EPA recommended methods. The dotted line is at the R-squared threshold 0.6. Other R-squared values in the cumulative distribution were for comparisons of other types of water quality measures, as indicated in the legend. The 48 data sets shown each have more than 30 paired samples.

Each row in Table C-1 corresponds to a symbol in Figure C-1. The rows are sorted in descending order based on the IA value. Note that the same methods can perform differently at different sites. This example highlights why application of site-specific alternative criteria is limited to the site where the water quality data were collected. The total number of paired samples is indicated by N. “Above LOQ” (n) is the number of paired samples that were above the LOQ for each method. The IA values are based on the paired samples that were above the LOQ.

Each row in Table C-2 corresponds to a symbol in Figure C-2. The rows are sorted in descending order based on the R-squared value. Note that the same methods can perform differently at different sites. This example highlights why application of site-specific alternative criteria is limited to the site where the water quality data were collected. The total number of paired samples is indicated by N. “Above LOQ” (n) is the number of paired samples that were above the LOQ for each method. The R-squared values are based on the paired samples that were above the LOQ. The abbreviations for Tables C-1 and C-2 are shown in Table C-3.

EPA chose the approach of using IA and R-squared to quantify the proportional and additive differences between two methods to assess their level of agreement. This was chosen over the approach of quantifying false positives and false negative rates achieved by the new, alternative method while considering the EPA-suggested method the gold standard. A false positive and false negative rate relies on the identification of a number for determining *true* and *false*. While error rates are one approach for evaluating the agreement of two methods, EPA believes that evaluating the additive and proportional differences provides more flexibility, and is more amendable to identifying alternative methods for site specific criteria.

EPA decided that an alternative indicator/method should involve a single assay and not a combination of different assays. This TSM does not support the use of multiple indicators or methods at the same time in a combinatory approach.

Table C-1. Data sets used (sorted in descending order of IA values)

Location	Contributor	Method 1	Method 2	IA (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay - Name
Phase II	Racine, WI	EC, MF	EC, MPN	0.97	98	85	MF-MPN same organism
Doheny 2007	SCCWRP	ENT, 1600	ENT, MPN	0.97	103	30	MF-MPN same organism
Doheny 2008	SCCWRP	ENT, 1600	ENT, MPN	0.95	337	148	MF-MPN same organism
Malibu 2009	SCCWRP	ENT, 1600	ENT, MPN	0.93	346	144	MF-MPN same organism
Doheny 2008	SCCWRP	ENT, 1600	ENT, qPCR	0.88	334	170	MF/MPN -qPCR same organism
Doheny 2008	SCCWRP	ENT, qPCR	ENT, qPCR UNC	0.81	334	206	qPCR-qPCR same order
Phase II	Racine, WI	ENT, MF	ENT, MPN	0.80	97	86	MF-MPN same organism
Malibu 2009	SCCWRP	ENT, 1600	ENT, IMS-ATP	0.79	209	112	other same organism
Little Venice	NOAA	ENT, MF	Human Bac, qPCR UCD	0.75	188	51	MF/MPN -qPCR different organism
Avalon 2008	SCCWRP	ENT, 1600	ENT, MPN	0.75	294	204	MF-MPN same organism
Ocean County sites	EPA Region 2	ENT, MF	ENT, qPCR	0.74	100	81	MF/MPN -qPCR same organism
Malibu 2009	SCCWRP	ENT, 1600	ENT, qPCR	0.73	337	250	MF/MPN -qPCR same organism
Hollywood-Broward; open water sites	NOAA	ENT, qPCR	Human Bac, qPCR UCD	0.73	194	108	qPCR-qPCR different organism
Avalon 2007	SCCWRP	ENT, 1600	ENT, MPN	0.72	240	191	MF-MPN same organism
Malibu 2009	SCCWRP	ENT, qPCR	ENT, qPCR UNC	0.70	246	173	qPCR-qPCR same order

Table C-1. Data sets used (sorted in descending order of IA values) (continued)

Location	Contributor	Method 1	Method 2	IA (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay - Name
Avalon 2008	SCCWRP	ENT, qPCR	EC, qPCR	0.69	197	189	qPCR-qPCR different organism
Little Venice	NOAA	ENT, MF	ENT, qPCR	0.65	188	151	MF/MPN -qPCR same organism
Beach study 1, North Beach, 36" depth	Racine, WI	EC, MPN	EC, qPCR UNC	0.64	64	34	MF/MPN -qPCR same organism
Beach study 2, North Beach	Racine, WI	EC, MPN	EC, qPCR UNC	0.63	140	32	MF/MPN -qPCR same organism
Hollywood-Broward; nearshore sites	NOAA	ENT, qPCR	Human Bac, qPCR UCD	0.63	64	45	qPCR-qPCR different organism
Little Venice	NOAA	ENT, qPCR	Human Bac, qPCR UCD	0.62	143	48	qPCR-qPCR different organism
Avalon 2008	SCCWRP	ENT, 1600	ENT, qPCR	0.60	245	213	MF/MPN -qPCR same organism
Doheny 2008	SCCWRP	ENT, 1600	GenBac, qPCR	0.60	334	217	MF/MPN -qPCR different organism
Beach study 2, Zoo Beach	Racine, WI	EC, MPN	EC, qPCR UNC	0.58	141	45	MF/MPN -qPCR same organism
Avalon 2008	SCCWRP	ENT, qPCR	ENT, qPCR UNC	0.58	159	132	qPCR-qPCR same order
Avalon 2008	SCCWRP	ENT, 1600	EC, qPCR	0.57	236	220	MF/MPN -qPCR different organism
Avalon 2007	SCCWRP	ENT, 1600	ENT, qPCR	0.57	306	231	MF/MPN -qPCR same organism
Doheny 2008	SCCWRP	ENT, qPCR	GenBac, qPCR	0.56	334	240	qPCR-qPCR different organism

Table C-1. Data sets used (sorted in descending order of IA values) (continued)

Location	Contributor	Method 1	Method 2	IA (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay - Name
Beach study 1, North Beach, 12" depth	Racine, WI	EC, MPN	EC, qPCR UNC	0.55	48	37	MF/MPN -qPCR same organism
Malibu 2009	SCCWRP	ENT, qPCR	GenBac, qPCR	0.53	337	300	qPCR-qPCR different organism
Beach study 2, North Beach	Racine, WI	ENT, MF	ENT, qPCR UNC	0.51	63	52	MF/MPN -qPCR same organism
Hawaii	Stanford	ENT, 1600	CP	0.46	88	80	culture-culture different organism
Malibu 2009	SCCWRP	ENT, 1600	GenBac, qPCR	0.46	337	268	MF/MPN -qPCR different organism
Hawaii	Stanford	ENT, qPCR	Human Bac, qPCR UCD	0.46	88	62	qPCR-qPCR different organism
Avalon 2008	SCCWRP	ENT, qPCR	GenBac, qPCR	0.46	245	230	qPCR-qPCR different organism
Hawaii	Stanford	ENT, qPCR	F+ phage	0.43	88	85	other different organism
Hawaii	Stanford	ENT, 1600	F+ phage	0.43	88	82	culture-culture different organism
Beach study 2, Zoo Beach	Racine, WI	ENT, MF	ENT, qPCR UNC	0.42	62	53	MF/MPN -qPCR same organism
Malibu (2009)	SCCWRP	ENT, 1600	Phage, 1601	0.42	161	58	culture-culture different organism
Hawaii	Stanford	ENT, 1600	ENT, qPCR	0.40	88	84	MF/MPN -QPCR same organism
Doheny (2007-2009)	SCCWRP	ENT, 1600	Phage, 1601	0.39	304	159	culture-culture different organism
Avalon 2008	SCCWRP	ENT, qPCR	CP, qPCR	0.39	219	205	QPCR-QPCR different organism
Hawaii	Stanford	EC, 1604	CP	0.36	87	82	culture-culture different organism

Table C-1. Data sets used (sorted in descending order of IA values) (continued)

Location	Contributor	Method 1	Method 2	IA (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay - Name
Hawaii	Stanford	EC, 1604	F+ phage	0.33	87	83	culture-culture different organism
Avalon 2008	SCCWRP	ENT, 1600	GenBac, qPCR	0.32	245	222	MF/MPN -qPCR different organism
Avalon (2007-2008)	SCCWRP	ENT, 1600	Phage, 1601	0.29	589	311	culture-culture different organism
Lake Erie	OSU	ENT, qPCR	EC, 1603	0.19	129	71	MF/MPN -qPCR different organism
Lake Erie	OSU	ENT, qPCR	ENT, IMS-ATP	0.15	83	44	other same organism

Table C-2. Data sets used (sorted in descending order of R-square values)

Location	Contributor	Method 1	Method 2	R-squared (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay Name
Doheny 2007	SCCWRP	ENT, 1600	ENT, MPN	0.91	103	30	MF-MPN same organism
Phase II	Racine, WI	EC, MF	EC, MPN	0.90	98	85	MF-MPN same organism
Doheny 2008	SCCWRP	ENT, 1600	ENT, MPN	0.84	337	148	MF-MPN same organism
Malibu 2009	SCCWRP	ENT, 1600	ENT, MPN	0.81	346	144	MF-MPN same organism
Phase II	Racine, WI	ENT, MF	ENT, MPN	0.65	97	86	MF-MPN same organism
Doheny 2008	SCCWRP	ENT, 1600	ENT, qPCR	0.63	334	170	MF/MPN -qPCR same organism
Hollywood-Broward; open water sites	NOAA	ENT, qPCR	Human Bac, qPCR UCD	0.59	194	108	qPCR-qPCR different organism
Doheny 2008	SCCWRP	ENT, qPCR	GenBac, qPCR	0.59	334	240	qPCR-qPCR different organism
Doheny 2008	SCCWRP	ENT, qPCR	ENT, qPCR UNC	0.57	334	206	qPCR-qPCR same order
Malibu 2009	SCCWRP	ENT, 1600	ENT, IMS-ATP	0.55	209	112	other same organism
Doheny 2008	SCCWRP	ENT, 1600	GenBac, qPCR	0.51	334	217	MF/MPN -qPCR different organism
Ocean County sites	EPA Region 2	ENT, MF	ENT, qPCR	0.44	100	81	MF/MPN -qPCR same organism
Little Venice	NOAA	ENT, qPCR	Human Bac, qPCR UCD	0.43	143	48	qPCR-qPCR different organism
Little Venice	NOAA	ENT, MF	Human Bac, qPCR UCD	0.41	188	51	MF/MPN -qPCR different organism
Avalon 2008	SCCWRP	ENT, 1600	EC, qPCR	0.39	236	220	MF/MPN -qPCR different organism
Avalon 2008	SCCWRP	ENT, 1600	ENT, MPN	0.39	294	204	MF-MPN same organism

Table C-2. Data sets used (sorted in descending order of R-square values) (continued)

Location	Contributor	Method 1	Method 2	R-squared (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay Name
Malibu 2009	SCCWRP	ENT, 1600	ENT, qPCR	0.36	337	250	MF/MPN qQPCR same organism
Beach study 1, North Beach, 36” depth	Racine, WI	EC, MPN	EC, qPCR UNC	0.35	64	34	MF/MPN -qPCR same organism
Malibu 2009	SCCWRP	ENT, qPCR	ENT, qPCR UNC	0.34	246	173	qPCR-qPCR same order
Little Venice	NOAA	ENT, MF	ENT, qPCR	0.34	188	151	MF/MPN -qPCR same organism
Avalon 2007	SCCWRP	ENT, 1600	ENT, MPN	0.32	240	191	MF-MPN same organism
Avalon 2008	SCCWRP	ENT, qPCR	EC, qPCR	0.31	197	189	qPCR-qPCR different organism
Beach study 1, North Beach, 12” depth	Racine, WI	EC, MPN	EC, qPCR UNC	0.29	48	37	MF/MPN -qPCR same organism
Avalon 2008	SCCWRP	ENT, 1600	ENT, qPCR	0.27	245	213	MF/MPN -qPCR same organism
Malibu 2009	SCCWRP	ENT, 1600	GenBac, qPCR	0.23	337	268	MF/MPN -qPCR different organism
Beach study 2, North Beach	Racine, WI	ENT, MF	ENT, qPCR UNC	0.23	63	52	MF/MPN -qPCR same organism
Malibu 2009	SCCWRP	ENT, qPCR	GenBac, qPCR	0.22	337	300	qPCR-qPCR different organism
Avalon 2008	SCCWRP	ENT, qPCR	ENT, qPCR UNC	0.17	159	132	qPCR-qPCR same order
Avalon 2008	SCCWRP	ENT, qPCR	GenBac, qPCR	0.15	245	230	qPCR-qPCR different organism
Doheny (2007-2009)	SCCWRP	ENT, 1600	Phage, 1601	0.15	304	159	culture-culture different organism
Hollywood-Broward; nearshore sites	NOAA	ENT, qPCR	Human Bac, qPCR UCD	0.15	64	45	qPCR-qPCR different organism

Table C-2. Data sets used (sorted in descending order of R-square values) (continued)

Location	Contributor	Method 1	Method 2	R-squared (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay Name
Hawaii	Stanford	ENT, 1600	CP	0.14	88	80	culture-culture different organism
Avalon 2007	SCCWRP	ENT, 1600	ENT, qPCR	0.14	306	231	MF/MPN -qPCR same organism
Beach study 2, North Beach	Racine, WI	EC, MPN	EC, qPCR UNC	0.11	140	32	MF/MPN -qPCR same organism
Beach study 2, Zoo Beach	Racine, WI	ENT, MF	ENT, qPCR UNC	0.10	62	53	MF/MPN -qPCR same organism
Lake Erie	OSU	ENT, qPCR	ENT, IMS-ATP	0.09	83	44	other same organism
Beach study 2, Zoo Beach	Racine, WI	EC, MPN	EC, qPCR UNC	0.09	141	45	MF/MPN -qPCR same organism
Avalon 2008	SCCWRP	ENT, qPCR	CP, qPCR	0.07	219	205	qPCR-qPCR different organism
Hawaii	Stanford	ENT, 1600	ENT, qPCR	0.05	88	84	MF/MPN -qPCR same organism
Hawaii	Stanford	EC, 1604	CP	0.05	87	82	culture-culture different organism
Hawaii	Stanford	ENT, 1600	F+ phage	0.03	88	82	culture-culture different organism
Avalon (2007-2008)	SCCWRP	ENT, 1600	Phage, 1601	0.03	589	311	culture-culture different organism
Lake Erie	OSU	ENT, qPCR	EC, 1603	0.02	129	71	MF/MPN -qPCR different organism
Avalon 2008	SCCWRP	ENT, 1600	GenBac, qPCR	0.02	245	222	MF/MPN -qPCR different organism
Hawaii	Stanford	ENT, qPCR	Human Bac, qPCR UCD	0.01	88	62	qPCR-qPCR different organism
Malibu (2009)	SCCWRP	ENT, 1600	Phage, 1601	0.01	161	58	culture-culture different organism

Table C-2. Data sets used (sorted in descending order of R-square values) (continued)

Location	Contributor	Method 1	Method 2	R-squared (BLOQ removed)	N	n (pairs both above LOQ)	Type of Assay Name
Hawaii	Stanford	EC, 1604	F+ phage	0.01	87	83	culture-culture different organism
Hawaii	Stanford	ENT, qPCR	F+ phage	0.00	88	85	other different organism

Table C-3. Methods abbreviations

Contributor	Abbreviation	Method Name or Citation
SCCWRP	ENT, 1600	EPA Method 1600 for enterococci (U.S. EPA, 2009a)
SCCWRP	ENT, qPCR	EPA Entero 1 (now EPA Method 1611) *Enterococcus* by qPCR (U.S. EPA, 2012b)
SCCWRP	ENT, MPN	IDEXX Enterolert
SCCWRP	ENT, qPCR UNC	Noble et al. (2010)
SCCWRP	EC, qPCR	*E. coli* EPA (Shanks) (Chern et al., 2011)
SCCWRP	GenBac, qPCR	*Bacteroides* EPA (Seifring et al., 2008)
SCCWRP	Phage, 1601	EPA Method 1601 (U.S. EPA, 2001a)
SCCWRP	Phage, 1602	EPA Method 1602 (U.S. EPA, 2001b)
SCCWRP	CP, qPCR	*C. perfringens* EPA (Shanks) (Lund et al., 2004)
SCCWRP	ENT, IMS-ATP	Lee et al. (2010)
Racine, WI	EC, MPN	*E. coli* IDEXX Colilert
Racine, WI	ENT, MF	Enterococci by MF – mEI EPA method 1600 (U.S. EPA, 2009a)
Racine, WI	EC, qPCR UNC	Noble et al. (2010)
Racine, WI	ENT, qPCR UNC	Noble et al. (2010)
NOAA	ENT, MF	EPA Method 1600 for enterococci (U.S. EPA, 2009a)
NOAA	ENT, MPN	IDEXX Enterolert
NOAA	ENT, qPCR	EPA Entero 1 (now EPA Method 1611) *Enterococcus* by qPCR (U.S. EPA, 2012b)
NOAA	Human Bac, qPCR UCD	Human-specific *Bacteroides* qPCR (Kildare et al., 2007)
OSU	ENT, qPCR	EPA Entero 1 (now EPA Method 1611) *Enterococcus* by qPCR (U.S. EPA, 2012b)
OSU	EC, 1603	*E. coli* EPA Method 1603 (U.S. EPA, 2009b)
OSU	EC, qPCR	*E. coli* EPA qPCR (Agidi et al., 2013; Lee and Deininger, 2004)
OSU	ENT, IMS-ATP	Immunomagnetic separation/adenosine triphosphate (New Horizon) (Agidi et al., 2013; Lee and Deininger, 2004)
Stanford	ENT, 1600	EPA Method 1600 for enterococci (U.S. EPA, 2009a)
Stanford	ENT, qPCR	Viau et al. (2011)
Stanford	EC, 1604	EPA Method 1604 for *E. coli* (U.S. EPA, 2002)
Stanford	CP	*C. perfringens* Hawaii Department of Health procedure (culture) (Viau et al., 2011)
Stanford	F+ phage	Boehm et al. (2009)
Stanford	Human Bac, qPCR UCD	Human-specific Bacteroidales qPCR (Viau et al., 2011)

Table C-3. Methods abbreviations (continued)

Contributor	Abbreviation	Method Name or Citation
EPA Region 2	ENT, MF	EPA Method 1600 for enterococci (U.S. EPA, 2009a)
EPA Region 2	ENT, qPCR	Haugland et al. (2005) (primers)

Note: EC = *E. coli*; MF = membrane filtration; MPN = most probable number; ENT = enterococci/*Enterococcus*; 1600 = EPA Method 1600; Human Bac = human *Bacteroides*; qPCR = quantitative polymerase chain reaction; UCD = University of California Davis; UNC = University of North Carolina; 1602 = EPA Method 1602; CP = *Clostridium perfringens*; 1603 = EPA Method 1603; 1604 = EPA Method 1604; IMS-ATP = immunomagnetic separation/adenosine triphosphate; IMS = immunomagnetic separation
SCCWRP = Southern California Coastal Water Research Project; NOAA = National Oceanic and Atmospheric Administration; OSU = Ohio State University; Stanford = Stanford University; WI = Wisconsin

Appendix D: Case Examples

Three examples using actual water quality data are presented in this appendix. The first two examples are from sites where the IA threshold is met (Sites A and B). The third example is from a site where neither IA nor the R-squared thresholds are met (Site C, Tropical Site). EPA chose these examples with a view to providing a selection of cases that differ in terms of threshold agreement and prevalence of data below the level of quantitation (LOQ).

Site A: Comparing Enterococci (culture) to Enterococci IMS-ATP

At Site A, 209 paired samples were evaluated by enterococci measured by culture and enterococci measured by the IMS-ATP method.

In this example, the raw data were first prepared by removing all paired observations for which either or both values were below the LOQ. A total of 97 paired observations were removed in this way. No data points were removed from the data set because they were above the upper limits of quantification. The data were then $\log_{10}$ transformed.

Figure D-1 is a graphical representation of the ENT (culture) and ENT (IMS-ATP) data for Site A after $\log_{10}$ transformation.

IA is calculated for the data sets using Equation 1 (same as shown in step 3 and Appendix C).

$$IA = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2} \qquad \text{[Equation 1]}$$

where x and y are the indicator/methods, *i* is a counter, *N* is the total number of data points in the data set, $\bar{x}$ and $\bar{y}$ are the averages of the x and y data sets, respectively, and *IA* varies from 0 to 1 (Willmott and Wicks, 1980) with 1 being perfect agreement. Here x represents the density of microorganisms determined by EPA indicator/method (EPA Method 1600) and y is the density of microorganisms determined by the alternative indicator/method (IMS-ATP).

The IA is 0.79, which is above the threshold of 0.7. The GM of 35 CFU per 100 mL would be the same for both enterococci measured by culture and enterococci measured by the IMS-ATP method. The STV would also be as described in the 2012 RWQC.

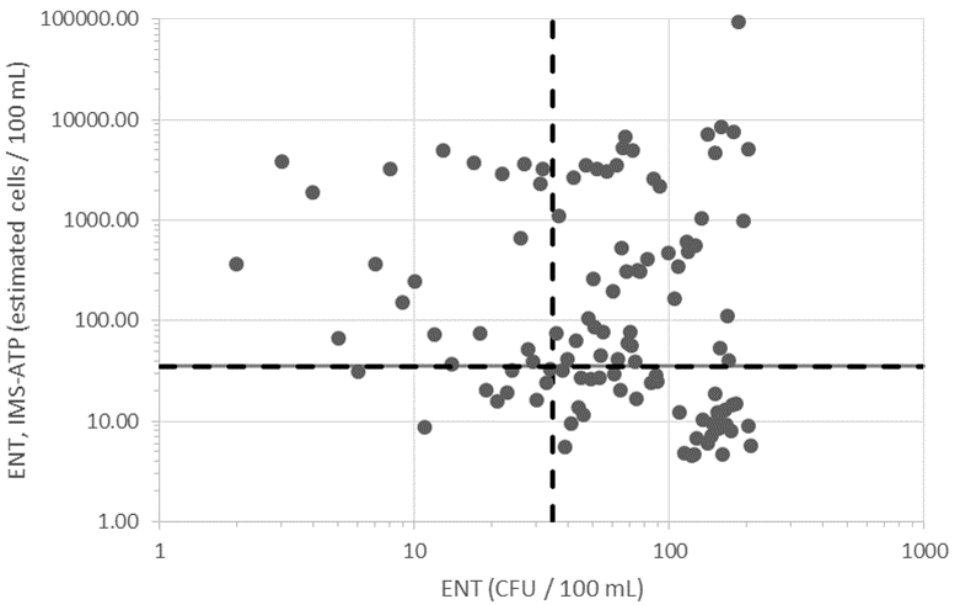

Figure D-1. Site A enterococci (culture) and enterococci IMS-ATP data presented graphically

Units for ENT-culture are CFUs. Units for enterococci measured using IMS-ATP are estimated cells. The vertical line is at 35 CFU per 100 mL and horizontal line is at 35 estimated cells per 100 mL.

<u>Site B: Comparing Enterococci (qPCR) to *Bacteroides* qPCR (BacHum)</u>

At Site B, 194 paired samples were evaluated by *Enterococcus* qPCR (ENT-qPCR)[26] and BacHum (human *Bacteroides* qPCR) methods.

In this example, the raw data were first prepared by removing all paired observations for which either or both values were below the LOQ. A total of 86 paired observations were removed in this way. No data points were removed from the data set because they were above the upper limits of quantification. The data were then $\log_{10}$ transformed.

Figure D-2 is a graphical representation of the ENT-qPCR and BacHum data for Site B after $\log_{10}$ transformation.

IA is calculated for the data sets using Equation 1 (same as shown in step 3 and Appendix C).

$$IA = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2} \qquad \text{[Equation 1]}$$

[26] In this case the ENT-qPCR method is similar to EPA Method 1611, with the exception that the standard curve was run with genomic DNA instead of calibrator cells.

where x and y are the indicator/methods, *i* is a counter, *N* is the total number of data points in the data set, $\bar{x}$ and $\bar{y}$ are the averages of the x and y data sets, respectively, and *IA* varies from 0 to 1 (Willmott and Wicks, 1980) with 1 being perfect agreement. Here x represents the density of microorganisms determined by EPA indicator/method (EPA Method 1611) and y is the density of microorganisms determined by the alternative indicator/method BacHum.

The IA is 0.73, which is above the threshold of 0.7. The GM of 300 CE per 100 mL would be the same for both ENT-qPCR and BacHum. The STV would also be as described in the 2012 RWQC.

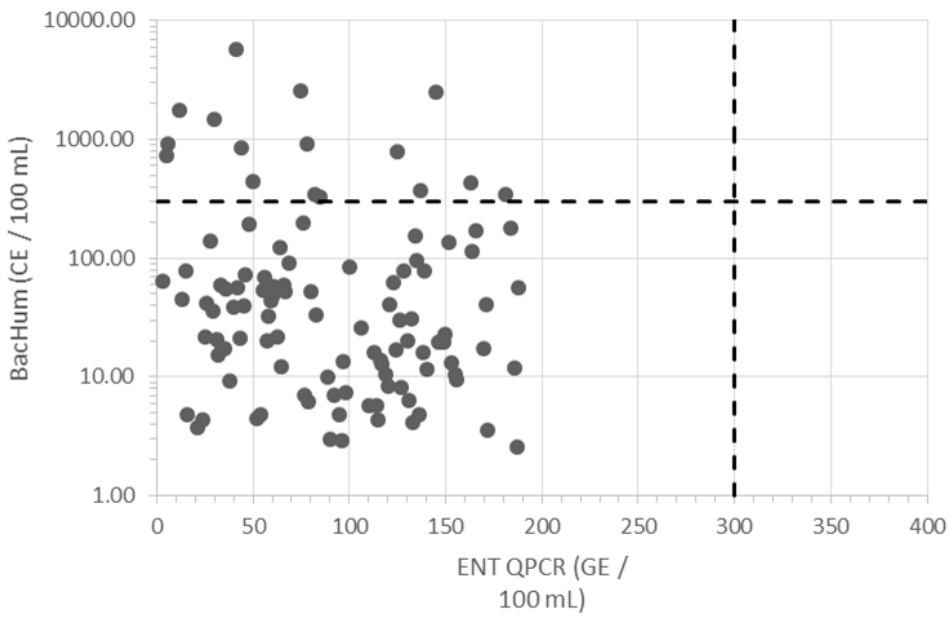

Figure D-2. Site B data presented graphically

Units for ENT-qPCR are genomic equivalents (GEs). The units for BacHum are cell equivalents (CE). For this example, we assume GEs are equivalent to CEs. The vertical and horizontal lines are at 300 CE per 100 mL.

<u>Site C: Comparing Enterococci (culture) to *C. perfringens*</u>

At this tropical site, 88 paired samples were evaluated for enterococci measured by culture and *C. perfringens* measured by culture.

In this example, the raw data were first prepared by removing all paired observations for which either or both values were below the LOQ. A total of 8 paired observations were removed in this way. No data points were removed from the data set because they were above the LOQ. The data were then $\log_{10}$ transformed.

Figure D-3 is a graphical representation of the enterococci and *C. perfringens* data for site C after $\log_{10}$ transformation.

The IA is calculated for the data sets using Equation 1 (same as shown in step 3 and Appendix C).

$$IA = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2} \qquad \text{[Equation 1]}$$

where x and y are the indicator/methods, *i* is a counter, *N* is the total number of data points in the data set, $\bar{x}$ and $\bar{y}$ are the averages of the x and y data sets, respectively, and *IA* varies from 0 to 1 (Willmott and Wicks, 1980) with 1 being perfect agreement. Here x represents density of microorganisms determined by the EPA indicator/method (EPA Method 1600) and y is the density of microorganisms determined by the alternative indicator/method *C. perfringens*. The IA is 0.46, which is below the threshold of 0.7.

In addition to the IA calculation, the case study example spreadsheet shows how the R-squared calculation is performed. The R-squared is 0.14, which is below the threshold of 0.6.

Both the IA and R-squared indicate that the *C. perfringens* method would not be well correlated with the enterococci method at this site. This is not surprising because tropical waters are known to harbor indigenous populations of enterococci that are not associated with human fecal contamination, whereas *C. perfringens* is associated with human fecal contamination and has fewer environmental sources than enterococci in tropical settings.

The approach presented in this TSM would not work for deriving site-specific alternative criteria for *C. perfringens* for this waterbody. In cases where this TSM does not work, the method may be a candidate for one of the other TSM approaches, such as *Site-Specific Alternative Criteria Technical Support Materials for Alternative Health Relationships.*

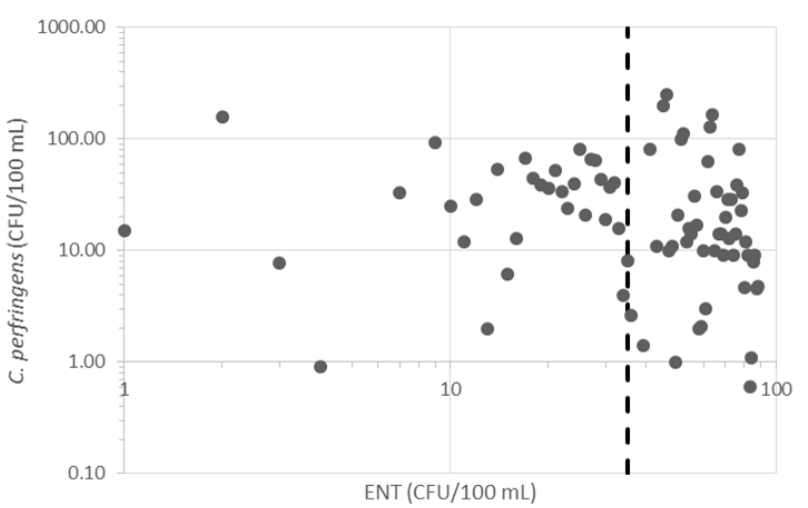

Figure D-3. Site C *C. perfringens* and enterococci data presented graphically

Units for enterococci and *C. perfringens* measured using culture are CFUs. The vertical line is at 35 CFU per 100 mL.

Appendix E: How to use Excel to calculate R-squared and index of agreement

Steps to setting up a spread sheet to calculate R-squared (RSQ) and Index of Agreement (IA).

Note: in the figures below, enterococci and clostridia results for only 8 samples are provided for purposes of illustration. This is not meant to imply that 8 samples are sufficient for this TSM.

Step 1. Paste your data into Excel. You should have two columns. Column A is the EPA method ("x"), Column B is the alternative method ("y"). In Figure E-1 below, you can see the top of the two columns for enterococci (ENT, the EPA method "x") and clostridia (the alternative method "y"). These data can be referred to as 'unscrubbed' because they may contain non-numerical entries like BLOQ (below limit of quantification).

ENT CFU/100 mL	clostridia CFU/100 m
unscrubbed	unscrubbed
x	y
290	15
8500	160
45	7.8
1	0.9
140	BLOQ
80	BLOQ
284	33
BLOQ	0.2

Figure E-1. Image of columns A and B (left to right)

Step 2. Create columns C and D which represent 'scrubbed' data from columns A and B, respectively. In these columns, non-numerical entries indicating below LOQ observations in either of the paired observations are replaced with blank entries. You can use an 'if statement' to do the replacement. The following if statement equation was used in column C to scrub the enterococci data in column A:

=IF(OR(A4="BLOQ",B4="BLOQ"),"",A4).

A similar statement could be used to create column D.

ENT CFU/100 mL	clostridia CFU/100 mL
scrubbed x	scrubbed y
290.00	15.00
8500.00	160.00
45.00	7.80
1.00	0.90
284.00	33.00

Figure E-2. Image of columns C and D (left to right)

Step 3. Create columns E and F which represent log_10 transformation of the scrubbed data (Figure E-3). In these columns, the log_10 of the data in columns C and D are calculated, respectively. You can use the function "LOG" in Excel, taking care to leave blank cells for which the corresponding non-transformed value is also blank. The following if statement equation was used in column E to log transform the data in column C:

=IF(C4="","",LOG(C4))

A similar statement could be used to create column F.

These are the columns that will be used to calculate the RSQ and IA.

ENT log (CFU/100 mL)	clostridia log (CFU/100 mL)
scrubbed log_10 x	scrubbed log_10 y
2.46	1.18
3.93	2.20
1.65	0.89
0.00	-0.05
2.45	1.52

Figure E-3. Image of columns E and F (left to right)

Step 4. Create columns G and H that will contain components of the numerator and denominator of the IA formula. Column G should contain the result of the following calculation: (logx_i-logy_i)^2. Column H should contain the results of the following calculation: (|logx_i-averagelogx|+|logy_i-averagelogx|)^2. Here logx_i and logy_i are the values from that row and averagelogx is the average of all the values in the logx_i column. This is how the code for Column G should look for the first data row:

=IF(OR(E4="",F4=""),"",(E4-F4)^2)

This is how the code from Column H should look in Excel for the first data row:

=IF(OR(E4="",F4=""),"",(ABS(E4-AVERAGE(E:E))+ABS(F4-AVERAGE(E:E)))^2)

These formulas will ignore any blank rows in the data carried across as a result of dropping BLOQ observations.

numerator in IA formula (logx_i-logy_i)^2	denominator in IA (|logx_i-averagex|+|y_i-average x|)^2
1.65	1.65
2.98	2.98
0.58	4.39
0.00	21.96
0.87	0.87

Figure E-4. Image of columns G and H (left to right)

Step 5. Calculate the R-squared using the following formula:

=RSQ(F:F,E:E)

This calculation can be done in an empty cell to the right of your columns (Figure E-5). The RSQ is calculated from the log_10 transformed data in columns E and F. This formula will ignore any blank rows in the data carried across as a result of dropping BLOQ observations.

Step 6. Calculate IA using this formula:

=1-(AVERAGE(G:G)/AVERAGE(H:H))

This calculation can be done in an empty cell to the right of your columns (Figure E-5). It uses the calculations in columns G and H. This formula will ignore any blank rows in the data carried across as a result of dropping BLOQ observations.

RSQ	**0.14**
IA	**0.46**

Figure E-5. Example of calculations for RSQ and IA

Appendix F: Example R Code

ANNOTATED CODE EXAMPLE: INDEX OF AGREEMENT AND R-SQUARED COMPUTATIONS

KEY
BOLD - Actual code
PLAIN – Comments for guidance
- Character to mark a comment in the R code

##

The following example code written in the R programming language (http://www.r-project.org/) facilitates the computation of the index of agreement and the R-squared metrics for a user-defined set of Y on X comparisons using log-transformed Y and X data.

To use this code, users must:

(i) provide the raw data for all water quality measures being compared in a specific format (discussed in Attachment 1 below)
(ii) provide a file defining the specific sets of Y on X comparisons to be performed in a specific format (discussed in Attachment 2 below) and
(iii) update the filepath and filenames corresponding to the raw data file and the comparisons file at the specified points in the code below.

Beginning of Code Example

Remove all residual variables from the workspace

rm(list=ls())

##

Note: USER INPUT REQUIRED HERE!

Reads in user-specified raw data filepath and name as a string. See required format in # Attachment 1 below.

#Specify between the inverted commas the filepath and filename of the raw data file #assembled in the required format

dbpathname<-"C:/filepath/datafile.csv"

##

##

Note: USER INPUT REQUIRED HERE: Specify filepath and filename containing desired Y on X #comparisons. See required format in Attachment 2 below.

```
comppathname<-"C:/filepath/comparisonsfile.csv"
```

Read in user-specified y on x comparisons filepath and filename as a string

```
##############################################
```

Clear previous output file if any

```
eval(parse(text=paste('unlink("',substr(dbpathname,1,nchar(dbpathname)-
4),'_output.csv",recursive=FALSE)',sep="")))
```

#Read in user-specified database into data frame "db"

```
eval(parse(text=paste('db<-
read.csv(file="',dbpathname,'",head=TRUE,sep=",",na.strings=c("NA", "NULL"))',sep="))) #
```

Store separately the 2nd and 3rd rows of the input dataset (containing BDL symbol and values) as dataframe "db_dl"

```
db_dl<-db[(1:2),]
```

Convert all variables to characters

```
db_dl <- data.frame(lapply(db_dl, as.character), stringsAsFactors=FALSE)
```

Drop the 2nd and 3rd rows of the input dataset (containing BDL symbol and values)

```
db<-db[-(1:2),]
```

Convert all variables to characters

```
db<- data.frame(lapply(db, as.character), stringsAsFactors=FALSE)
```

Loop to replace BDL/ND with user-specified replacement values

```
for (i in 1:ncol(db)) {
db[,i]<-ifelse (db[,i]==db_dl[1,i],db_dl[2,i],db[,i])
}
```

Convert all variables to numeric.

```
db<- data.frame(lapply(db, as.numeric), stringsAsFactors=FALSE)
```

Define goodness-of-fit function “gof”. Arguments are x variable name and y variable name

```
gof<-function(x,y){ db_f<-db[c(x,y)]
```

Keep only x and y variables in data frame "db_f"

```
# Delete observations with missing values from "db_f"
db_f <- na.omit(db_f)
# Rename input x variable as x
eval(parse(text=paste('names(db_f)[names(db_f)=="',x,'"]<-"x"',sep="")))
# Rename input y variable as y
eval(parse(text=paste('names(db_f)[names(db_f)=="',y,'"]<-"y"',sep="")))
# LOG ANALYSIS
# Introduce column with logarithm 10 of x
db_f$lx<-log10(db_f$x)
# Introduce column with logarithm 10 of y
db_f$ly<-log10(db_f$y)
# Delete observations with missing values
db_f <- na.omit(db_f)
# Regress ly on lx. The output of this regression will provide the R-squared  value
regl<-lm(ly~lx,data=db_f)
# Mean of lx
lxm<-mean(db_f$lx)
# Mean of ly
lym<-mean(db_f$ly)
# Calculate square error (on log values)
db_f$lsqer<-(db_f$lx-db_f$ly)^2
# Calculate mean square error(on log scale)
lmse<-mean(db_f$lsqer)
# Calculate square absolute error (with respect to x)  (on log values)
db_f$lsqabserr<-(abs(db_f$lx-lxm)+abs(db_f$ly-lxm))^2
# Calculate mean of square of absolute error (on log values)
```

```
lmsqabserr<-mean(db_f$lsqabserr)

# Calculate Index of Agreement (on log values)

lioa<-1-lmse/lmsqabserr

# Output vectors for log analysis

o1<-cbind(paste(y,"v/s",x,sep=" "),"Log-Transformed Data","R-
squared",summary(regl)$r.squared)

o2<-cbind(paste(y,"v/s",x,sep=" "),"Log-Transformed Data","Index of Agreement",lioa)

o3<-cbind(paste(y,"v/s",x,sep=" "),"Log-Transformed Data","N",nrow(db_f))

# Bind and write all output vectors as .csv file

# Column bind output vectors into composite output vector opvec

opvec<-cbind(o1,o2,o3)

# Outputs results as a .csv file

eval(parse(text=paste('write(opvec,file="',paste(substr(dbpathname,1,nchar(dbpathname)-
4),"_output.csv",sep=""),'",ncolumns=4,append =TRUE, sep = ",")',sep="")))
}

# End of Function GOF

# Reads in user-specified comparison database into data frame "comp"

eval(parse(text=paste('comp<-
read.csv(file="',comppathname,'",head=TRUE,sep=",",na.strings=c("NA", "NULL"))',sep="")))

# Convert all variables to character

comp <- data.frame(lapply(comp, as.character), stringsAsFactors=FALSE)

# For loop to process Y on X comparisons defined in comparisons file

for (j in 1:nrow(comp)){
eval(parse(text=paste('gof("',comp[j,1],'","',comp[j,2],'")',sep="")))
}

# End of Code Example

# Note: The output of this procedure will be stored in a .csv file in the same location as the raw
data file. The output file name will be identical to the raw data filename, but suffixed with a
"_output" tag. The format of the output file is explained in Attachment 3.
```

The R code with no annotation is in the text box below. It can be copied and pasted directly into R.

```
# Code Start


rm(list=ls())

##############################################

# USER INPUT REQUIRED HERE!

dbpathname<-"C:/Filepath/FilenameofInputData.csv"

##############################################


##############################################

# USER INPUT REQUIRED HERE!

comppathname<-"C:/Filepath/FilenameofComparisonsFile.csv"

##############################################


eval(parse(text=paste('unlink('',substr(dbpathname,1,nchar(dbpathname)-
4),'_output.csv",recursive=FALSE)',sep="")))

eval(parse(text=paste('db<-read.csv(file='',dbpathname,'',head=TRUE,sep=",",na.strings=c("NA",
"NULL"))',sep="")))

db_dl<-db[(1:2),]

db_dl <- data.frame(lapply(db_dl, as.character), stringsAsFactors=FALSE)

db<-db[-(1:2),]

db<- data.frame(lapply(db, as.character), stringsAsFactors=FALSE)

for (i in 1:ncol(db)){

db[,i]<-ifelse (db[,i]==db_dl[1,i],db_dl[2,i],db[,i])

}

db<- data.frame(lapply(db, as.numeric), stringsAsFactors=FALSE)

gof<-function(x,y){
```

```
db_f<-db[c(x,y)]

db_f <- na.omit(db_f)

eval(parse(text=paste('names(db_f)[names(db_f)=="',x,'"]<-"x"',sep="")))

eval(parse(text=paste('names(db_f)[names(db_f)=="',y,'"]<-"y"',sep="")))


db_f$lx<-log10(db_f$x)

db_f$ly<-log10(db_f$y)

db_f <- na.omit(db_f)

regl<-lm(ly~lx,data=db_f)

lxm<-mean(db_f$lx) # Mean of lx

lym<-mean(db_f$ly) # Mean of ly


db_f$lsqer<-(db_f$lx-db_f$ly)^2

lmse<-mean(db_f$lsqer)

db_f$lsqabserr<-(abs(db_f$lx-lxm)+abs(db_f$ly-lxm))^2

lmsqabserr<-mean(db_f$lsqabserr)

lioa<-1-lmse/lmsqabserr

o1<-cbind(paste(y,"v/s",x,sep=" "),"Log-Transformed Data","R-squared",summary(regl)$r.squared)

o2<-cbind(paste(y,"v/s",x,sep=" "),"Log-Transformed Data","Index of Agreement",lioa)

o3<-cbind(paste(y,"v/s",x,sep=" "),"Log-Transformed Data","N",nrow(db_f))

opvec<-cbind(o1,o2,o3)

eval(parse(text=paste('write(opvec,file="',paste(substr(dbpathname,1,nchar(dbpathname)-
4),"_output.csv",sep=""),'",ncolumns=4,append =TRUE, sep = ",")',sep="")))

}


eval(parse(text=paste('comp<-read.csv(file="',comppathname,'",head=TRUE,sep=",",na.strings=c("NA",
"NULL"))',sep="")))

comp <- data.frame(lapply(comp, as.character), stringsAsFactors=FALSE)

for (j in 1:nrow(comp)){
```

```
eval(parse(text=paste('gof("',comp[j,1],'","',comp[j,2],'")',sep="")))

}

# Code End
```

Attachment 1: Required Format for Raw Data .csv File

- The first row of the raw data file should contain the variable names for the sampled water quality indicators/pathogens, each occupying its own column in the .csv file (comma-separated values file). Note: csv files may be created within MS-Excel by saving as ".csv".
- The second row must contain the below detection level or below level of quantitation symbol used for each water quality indicator/pathogen.
- The third row should contain the replacement value that should be assigned for below detection level and below LOQ data. Given the recommended approach of dropping below LOQ data, this row should be left blank.
- The fourth row onwards contains sampled water quality indicator/pathogen values. Note that each row should contain contemporaneously sampled data; in other words, all the data within a given row should have been obtained at an identical point in time. When the measured value is below the detection limit or below the LOQ, the appropriate symbol defined in the second row should be used.
- The data must be reported in original units and not log transformed. The program will perform the log transformation.

Table F-1 provides an example of an acceptable data format for the raw data .csv file.

Table F-1: Example of raw data .csv file (required input for the program)

enterococci	E.coli	C.perfringens	F.Phage	EnterococcusQ	Bachum
BDL	BDL	BDL	BDL	ND	ND
290	330	15	10.8	287.1862	814.4115
8500	3400	160	2.6	143.6567	769.3569
45	220	7.8	5	0.9064	137.0487
1	990	0.9	1	3.0738	72.1189
140	190	BDL	4.8	942.0926	2.335
80	300	BDL	39	5.3288	46.1044

BDL = below detection level

Attachment 2: Required format for comparisons .csv file

- The first row, first column must contain the letter X.
- The first row, second column must contain the letter Y.
- The subsequent rows should contain the X-Y comparisons to be performed specified in terms of the same variable names as in the raw data file.
- It is critical to specify X and Y variables in the appropriate place. The Y variable should be imagined to appear on the Y axis of a graph, and the X variable on the X axis. This is referred to as a Y on X comparison in which Y is the dependent (or predicted) variable and X is the independent (or predictor) variable. Note that the results of a Y on X comparison are different from those of an X on Y comparison.
- For example, to predict “Ecoli” levels based on “Enterococci” data, put “Ecoli” under Y and “Enterococci” under X.
- It is recommended that the Comparisons file be stored in the same filepath location as the raw data file for clarity and later reference.

Table F-2 provides an example of a comparisons file based on the variables specified in the raw data file example above.

Table F-2: Example of input comparisons .csv file (required input for the program)

X	Y
Enterococci	EnterococcusQ
Enterococci	F.Phage
EnterococcusQ	Bachum
EnterococcusQ	F.Phage
F.Phage	C.perfringens
Bachum	F.Phage

Attachment 3: Format of the Output csv File

- The output file is a .csv file.
- The output file will be stored in the same location as your raw data file.
- The output filename will be identical to your raw data file name, except that it will be suffixed by the term "_output".
- The first column in the output file describes the comparison that was performed as "Y v/s X". This means that Y is the dependent or predicted variable and X is the independent or predictor variable.
- The second column in the output file indicates the operation was performed on log-transformed data.
- The third column describes the metric that was computed for the variables being compared. This includes R-squared, Index of Agreement and the number of available observations (N).
- The fourth column contains the corresponding value or score for the computed metric.

Table F-3 provides an illustrative example of an output file based on the variables and comparisons specified in Tables F-1 and F-2 above. Not that the output values in Table F-3 do not correspond to the data presented in Table F-1 but are based on a larger dataset.

Table F-3: Example of output .csv file (sample output from the program)

Column 1	Column 2	Column 3	Column 4
EnterococcusQ v/s enterococci	Log-Transformed Data	R-squared	0.052829
EnterococcusQ v/s enterococci	Log-Transformed Data	Index of Agreement	0.431558
EnterococcusQ v/s enterococci	Log-Transformed Data	N	88
F.Phage v/s enterococci	Log-Transformed Data	R-squared	0.060856
F.Phage v/s enterococci	Log-Transformed Data	Index of Agreement	0.467949
F.Phage v/s enterococci	Log-Transformed Data	N	88
bachum v/s EnterococcusQ	Log-Transformed Data	R-squared	0.033564
bachum v/s EnterococcusQ	Log-Transformed Data	Index of Agreement	0.510555
bachum v/s EnterococcusQ	Log-Transformed Data	N	88
F.Phage v/s EnterococcusQ	Log-Transformed Data	R-squared	0.000622
F.Phage v/s EnterococcusQ	Log-Transformed Data	Index of Agreement	0.381447
F.Phage v/s EnterococcusQ	Log-Transformed Data	N	88
C.perfringens v/s F.Phage	Log-Transformed Data	R-squared	0.052369
C.perfringens v/s F.Phage	Log-Transformed Data	Index of Agreement	0.508983
C.perfringens v/s F.Phage	Log-Transformed Data	N	88
F.Phage v/s bachum	Log-Transformed Data	R-squared	0.011579
F.Phage v/s bachum	Log-Transformed Data	Index of Agreement	0.319352
F.Phage v/s bachum	Log-Transformed Data	N	88
F.Phage v/s E.coli	Log-Transformed Data	R-squared	0.019159
F.Phage v/s E.coli	Log-Transformed Data	Index of Agreement	0.336693
F.Phage v/s E.coli	Log-Transformed Data	N	87
C.perfringens v/s E.coli	Log-Transformed Data	R-squared	0.014895
C.perfringens v/s E.coli	Log-Transformed Data	Index of Agreement	0.339709
C.perfringens v/s E.coli	Log-Transformed Data	N	87