EVALUATION OF MULTIPLE INDICATOR COMBINATIONS TO DEVELOP QUANTIFIABLE RELATIONSHIPS

U.S. Environmental Protection Agency

Office of Water
December 2010
EPA 822-R-10-004

Table of Contents

List of Tables

List of Figures

Executive Summary

This report was written to meet the EPA Critical Path Science Plan[1] element P15. The objectives of this work were to compare the fecal indicator bacteria and health effect relationships for multiple indicator/method combinations used in epidemiology studies and to evaluate multiple indicator/method combinations to develop quantifiable relationships. To meet those objectives, two approaches were employed to relate the different indicator/method combinations to gastrointestinal (GI) illness risks, namely the Risk Link and Water Quality Link.

In the Risk Link approach, indicator-method combinations are linked via health effects curves[2] generated by epidemiology studies. Three demonstrations of the Risk Link approach are presented. In the first demonstration, linkages are established between indicator densities for multiple indicators used in the same epidemiology studies. This demonstration is a straightforward implementation of the Risk Link approach and entails linkage of *Enterococcus* density as measured by qPCR with *Bacteroidales* density as measured by qPCR. The second demonstration illustrates an assessment of differences between epidemiology studies conducted at different places or times, using an analysis of covariance (ANCOVA) to compare data from (Marion et al.[3]) and (USEPA[4]) EPA recreational water epidemiology studies in freshwater. The analyses indicate that although the two sets of studies were conducted at different times, the slopes of their health effects curves are statistically similar.

The third Risk Link was performed using health effects relations from epidemiology studies with *Enterococcus* and GI relationships, measured by either qPCR or culture-based method. The EPA recreational water epidemiology study in marine waters[5] and the EPA's National Epidemiological and Environmental Assessment of Recreational (NEEAR) Water Studies (marine) were linked to determine comparable *Enterococcus* culture and qPCR indicator densities at various GI health risk levels. To enable the calculation of comparable *Enterococcus* densities for the two studies, results from the 1983 USEPA study were first translated such that the GI illness definition matched that of the NEEAR GI illness (NGI) definition. Although numeric qPCR-based water quality criteria have not been established at this time, assumptions about the way those criteria will likely be developed can be used to demonstrate potential qPCR and culture-based criteria that are consistent with the same level of risk (table below).

[1] USEPA. 2007. *Critical Path Science Plan for the Development of New or Revised Recreational Water Quality Criteria*. U.S. Environmental Protection Agency, Offices of Water and Research and Development, Washington, DC.

[2] A health effects curve refers to a mathematical relationship between fecal indicator bacteria (FIB) density and observed illness in epidemiology studies of recreational waters. All health effects curves referenced in this report relate the incidence of gastrointestinal (GI) illness to the log-transformed FIB density

[3] Marion, J.W., Lee, J., Lemeshow, S., Buckley, T.J. 2010. Association of illness and recreational water exposure during advisory and non-advisory conditions at an inland U.S. beach. Water Research 44(16): 4796-4804.

[4] USEPA 1984. *Health Effects Criteria for Fresh Recreational Water.* EPA-600/1-84-004. Research Triangle Park, NC: USEPA.

[5] USEPA 1983. *Health Effects Criteria for Marine Recreational Waters*. EPA-600/1-80-031. Research Triangle Park, NC: USEPA.

Tolerable attributable illness level (as HCGI‡ per 1000 swimmers)	Tolerable attributable illness level (as NGI* per 1000 swimmers)	Hypothetical qPCR *Enterococcus* density (CCE§/100 mL)	Hypothetical geometric mean membrane filtration *Enterococcus* density (CFU/100 mL)	75th percentile value† (CFU/100 mL)
8	35	427	11	33
10	43	610	14	42
19	82	3460	35	104

† The 75th percentile value for *Enterococcus* density based on the calculated geometric mean *Enterococcus* density, assuming *Enterococcus* densities are log-normally distributed, and assuming a typical standard deviation of log-transformed *Enterococcus* density for marine sites of 0.7.

‡ Highly Credible Gastrointestinal Illness, as defined in the epidemiology studies conducted in support of the 1986 water quality criteria

* NEEAR study GI illness (NGI, per the definition used in the NEEAR epidemiology studies)

§ Calibration cell equivalent; using a calibration sample containing a known concentration of the target sequence, CCEs are the normalized values of the test sample cell equivalents

The Water Quality Link approach links an indicator-method combination for which there is no health effects relation to an indicator-method combination with a health effects relation via a quantifiable relationship between their measured fecal indicator bacteria densities. In short, the Water Quality Link Approach uses paired water quality data, rather than the linkage of health effects curves alone, as the basis for an alternative method for establishing culture-based criteria. Demonstrations of this approach were made using densities of *Enterococcus* enumerated by qPCR and by membrane filtration (MF) and linked to the health effects relationship from EPA's NEEAR freshwater studies. It was found that while the Water Quality Link may be useful on a site specific basis, in this particular demonstration the relationships between paired *Enterococcus* data, measured by culture and qPCR, were not consistent among NEEAR study freshwater beaches when simple linear and broken stick (segmented) regression models were employed. Further, the regression fits to this dataset exhibited heteroskedasticity (uneven distribution of residuals) and both the simple linear regression and broken-stick models resulted in comparable *Enterococcus* culture criteria values far in excess of the current criteria.

The exploration of the Risk Link and the Water Quality Link approaches provides a proof of concept with the currently available data. It is likely that new health effects data or improved models of the co-occurrence of indicators will allow for additional or improved linkages to be established.

1 Introduction

1.1 Purpose, Objectives, and Importance

Pathogens within fecal matter arising from human and animal sources that enter coastal and inland waters pose risks to recreational swimmers. Because it is not feasible to monitor for all pathogens that may occur in ambient waters, monitoring strategies currently involve the detection of indicator organisms that are present in fecal material in greater numbers than the pathogenic organisms (NRC 2004). For decades, EPA has relied on the use of epidemiology studies to assess the risk of gastrointestinal (GI) illness in swimmers exposed to increasing densities of fecal indicator bacteria (FIB).[6] Previous epidemiology studies used only culture-based methods for enumerating FIB. More recent epidemiology studies have evaluated associations between GI illness and a wider range of indicator organisms, using both improved culture techniques and molecular-based (rapid) methods.

The primary purpose of this report is to help meet one of the elements (Project P15) in the U.S. Environmental Protection Agency (EPA or the Agency) *Critical Path Science Plan for Development of New or Revised Recreational Water Quality Criteria* (CPSP or science plan) (USEPA 2007). The science plan is a key component of EPA's overall process to develop new or revised Section 304(a) Ambient Water Quality Criteria (RWQC) for recreational waters that will be used by States, Territories, and Tribes to develop their own water quality standards (WQS[7]).

The objectives of CPSP Project P15 are as follows:

- To compare the GI illness response to exposure relationship for multiple fecal indicator organism/method combinations; and
- To develop quantifiable relationships between the results from the various indicator/method combinations.

More generally, P15 commissions scientific studies to explore correlations and linkages between available indicator organisms (FIB) and methods. These linkages are established to ensure that equal protection is provided for any future criteria associated with each indicator-method combination.

Section 2 provides a background review and assessment of relevant epidemiology studies, as well as a review of differences in FIB performance for indicators enumerated culture and quantitative polymerase chain reaction (qPCR). The background supports the analyses described in Section 3 that analyze available epidemiology datasets using both the "Risk Link" and the

[6] "Traditional" FIB include culturable total coliforms, fecal (thermotolerant) coliforms, *Escherichia coli* (an important member of the coliform group), and *Enterococcus* (enterococci). Although the presence of FIB indicates the presence of fecal matter, and the potential presence of (enteric) pathogens, FIB are not pathogenic (NRC 2004; WHO 2003).

[7] Under the Clean Water Act (CWA) States, Territories, and Tribes are required to adopt new or revised WQS for those pathogens and pathogen indicators for which EPA's new or revised criteria have been developed. Once approved by EPA, WQS are used for various CWA purposes and programs and are the effective (enforceable) standards to protect waters for specified designated uses, such as "primary contact recreation."

"Water Quality Link" approaches. Both of these approaches are described in detail in later sections of this report. Section 4 provides a discussion of the analyses, including limitations and assumptions.

1.2 Approaches for Developing Water Quality Measures Corresponding to Equivalent Risks

At the core of RWQC is a quantitative relationship between a measure of fecal pollution in water and the risk of adverse health outcomes (health effects curves) arising from primary recreational contact (e.g., swimming) in the polluted water. Although the expense, complexities, and time associated with conducting epidemiology studies are significant, several studies have been completed and are ongoing in the U.S. and abroad since the issuance of the current (1986) RWQC. Many of those studies were conducted in direct support of the development of RWQC, standards, and guidance. In addition, new technologies, such as rapid molecular-based detection methods, have been developed for identifying and enumerating FIB and other (alternate) indicator organisms that are fundamentally different from traditional culture-based methods. None of these approaches (new methods, alternative indicators, or alternate risk assessment methods) involves direct measurement of the pathogen(s) suspected of causing illness in recreational water users. Therefore, linkages are made between the indicators that are being measured and the adverse health effects observed in relevant and available epidemiology studies.

The health effects curves from some of the epidemiology studies may not be compared without prior analysis and transformations. Differences that necessitate transformations include, for example, differences in study design (prospective cohort [PC] vs. randomized control trial ([RCT]) (see more below), differences in definitions of GI illness, and differences in the settings (marine vs. freshwater environments). Ongoing epidemiology studies may yield health effects curves for additional indicator/method combinations and these health curves may be suitable for use in developing linkages between indicator/method combinations.

Further, incorporation of molecular-based methods into RWQC based on culture methods (or incorporation of culturable methods into criteria based on molecular methods) is complicated by the differences in their targets. Because different methods measure different targets with different abundances in environmental samples, it is not a given that health effect relationships developed for one particular indicator-method combination are valid for another indicator-method combination. This report places emphasis on the linkages between (1) *Enterococcus* measured by qPCR and *Bacteroidales* measured by qPCR for publicly owned (sewage) treatment works (POTW)-impacted marine settings; (2) *E. coli* measured by membrane filtration (MF) for studies conducted at different times and in different settings; and (3) *Enterococcus* measured by qPCR and *Enterococcus* measured by MF for studies of marine beaches conducted at different times.

1.2.1 Risk Link

1.2.1.1 Approach Description

The most direct route to meeting the CPSP P15 objectives is through use of health effect-FIB density curves to determine quantifiable relationships for indicator/method combinations that correspond to the same risk level for swimmers. This approach, termed the Risk Link, is illustrated in Figure 1. For this approach to be feasible, all indicator-method combinations

should have an associated relationship between indicator density and health effects (the solid lines in both plots in Figure 1). At present, all such health effects curves have been developed based on epidemiology studies of surface water recreation sites. Future relationships may be developed using QMRA, watershed modeling approaches, or other methods developed specifically for that purpose. Assuming the indicator densities and illness rates for the standard and alternative indicator health curves are comparable (and this assumption is addressed below), the indicator density corresponding to a selected acceptable level of risk is calculated using the health effects relationship for each of the indicator/illness combinations.

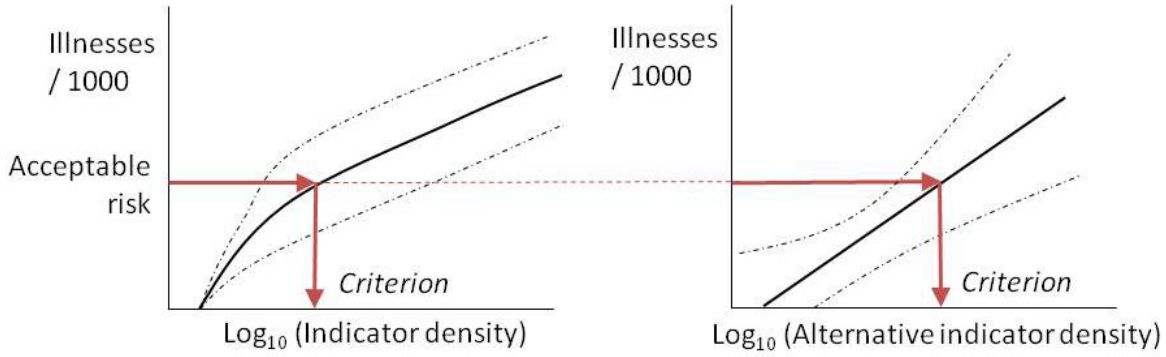

Figure 1. Illustration of the risk link approach

The rationale for using a Risk Link approach is that it maximizes use of direct measurements of health effects from relevant and available epidemiology studies. Over the recreational season, the pathogens to which swimmers are exposed and the densities of fecal indicators are variable. Further, the characteristic pathogens and FIB densities at different beaches likely differ, depending on the alignment of the beach with fecal pollution sources and other factors such as climate and rainfall. Given this variability, measurement of the association of adverse health effects with FIB through epidemiology studies appears to be the most direct and reliable means for relating water quality to health effects.

1.2.1.2 Harmonizing Data from Disparate Studies

For health effects curves to be comparable, they should relate to risks for the same illness. Further, the statistical measure used for characterizing water quality (e.g., geometric mean of multiple samples on a single day vs. the water quality from a single sample taken in a zone and at a time where swimming occurs) should be taken into account. An extreme example of curves that would not be comparable is curves that correspond to excess GI illness in swimmers and excess respiratory infection in non-swimmers. A summary of potential differences between epidemiology studies that could prevent direct comparison of health effects curves is presented in Table 1.

Table 1. Epidemiology study differences that limit direct comparison of health effects curves

Cause	Differences in data construct
Different epidemiology study design (PC vs. RCT)	• Exposures in prospective cohort (PC) studies are not prescribed and have a much wider range of durations and ingestions typical of recreation events; exposures in randomized control trial (RCT) studies are controlled and have more consistent ingestion volume, but less variability in exposure. • RCT studies control the location of exposure far more than in PC studies; thus, the impact of this element of study design is site specific. • Water quality associated with illness incidence in PC studies is based on an average indicator (usually FIB) density for the recreation site and over the entire study day; water quality associated with illness incidence in RCT studies is based on the indicator density for a sample taken at the same time and location as swimmer exposure. Proponents of RCT designs state that this feature reduces bias, but this claim may not be true depending on the magnitude of short-term variability.
Different definition of human health outcome	Epidemiology studies differ in their choices in health endpoints and their definitions of those endpoints, including GI illness.
Different settings	Indicators may be associated with different risks at sites with different • fecal pollution sources; • level of treatment of fecal pollution; • loading characteristics (continuous vs. event); and • proportion of FIB resuspended in sediments. Examples of pairs of studies that may have significant setting-related effects include • studies conducted in inland and coastal waters, • studies conducted in the United States and Europe, and • studies conducted at sites in (sub)tropical and temperate climates.
Regulation changes and technological advances	Regulations and advances in technology have changed wastewater (including POTW) treatment practices and the resulting indicator and pathogen loads to recreation sites. The most important of these changes are disinfection of wastewater, development of improved animal waste treatment technologies, and regulations on the land application of biosolids (treated wastewater sludge) and animal wastes.

1.2.2 Water Quality Link

Many indicator/method combinations are not, at present, associated with health effects curves or have not been applied directly in epidemiology studies. In such cases, a less direct approach for linking indicator/method combinations can be applied. The Water Quality Linkage approach relates paired data from two indicator/method combinations and links the method/indicator combinations to each other through their relationship, rather than through a direct linkage to health effects curves. This approach is illustrated in Figure 2, where in the top graph, a criterion density for a standard indicator/method combination is established based on a selected level of tolerable risk and a health effects curve. A quantifiable relationship for an alternative method/indicator combination is established through a model relating the density of the standard indicator/method combination to the density of the alternative method/indicator combination.

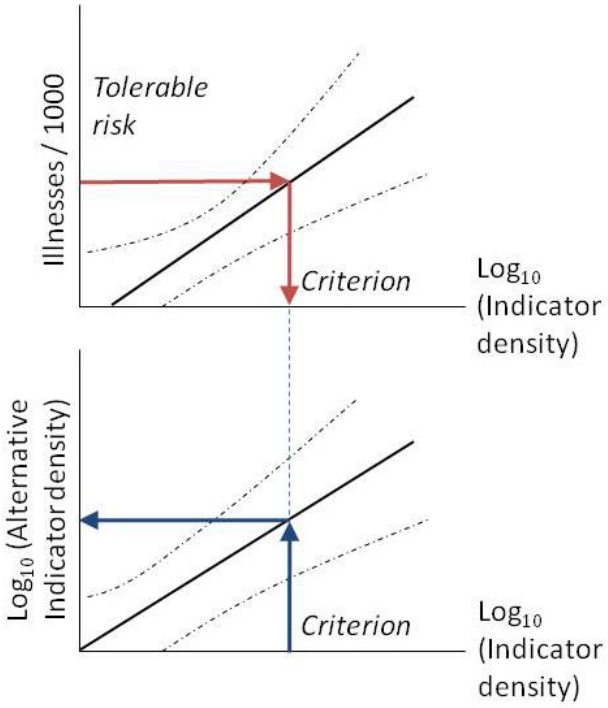

Figure 2. Illustration of the Water Quality Linkage approach

As with the Risk Link approach, the Water Quality Link should be employed when paired water quality data are from comparable studies. Further, a Water Quality Linkage at one site may not be applicable to another site—even if the sites have the same primary fecal pollution impact. A critical component in the application of the Water Quality Link is the development of a useful statistical model that relates FIB-method combinations. At present, models relating culturable and qPCR indicator counts are either site-specific (e.g., Byappanahalli et al. 2010; Lavender and Kinzelman 2009) or have been developed based on pooling of datasets under the assumption that datasets are similar and may be pooled (Haugland et al. 2005; Whitman et al. 2010). In this report, linear, and "broken stick" models for relating log-transformed culture and qPCR indicator densities are explored. It is possible that other statistical models linking these types of analytical methods will be proposed and evaluated in the future.

2 Background

Ordinarily, FIB themselves do not cause illness, but their densities can provide estimates overall levels of fecal contamination. As a result, their densities cannot usually be used to directly estimate health risks through a risk assessment approach. The relationship between indicators and health risks are best and most commonly established by epidemiology studies (NRC 2004; WERF 2009). Epidemiology studies (1) establish microbial water quality, typically via FIB density measurements, with a sufficient number and timing of samples to establish a characteristic indicator density during swimming; and (2) associate the FIB density with the adverse health effects observed among the population swimming compared to a non-swimming control population. This section includes a review and assessment of relevant and available epidemiology studies, as well as a review of differences in FIB performance for culture and qPCR indicators.

2.1 Review and Assessment of Relevant Epidemiology Studies and Datasets

Since the 1950s, numerous epidemiology studies have been conducted in the United States and abroad, most commonly at beaches impacted by sewage/wastewater effluent (e.g., POTWs), to evaluate the association between recreational water quality and adverse health outcomes. In these studies, attempts were made to relate a quantitative microbial indicator of water quality to health effects (usually some form of GI illness) using log-transformed data or a geometric mean to characterize exposure to water quality indicators. However, eye infections; skin irritations; ear, nose, and throat infections; and respiratory illness have also been evaluated. Many of these recreational water epidemiology studies are reviewed in one or more of the meta-analyses/systematic reviews of Prüss (1998), Wade et al. (2003), and Zmirou et al. (2003). More recent studies are reviewed in WERF (2009), which notes that all recreational epidemiology studies identified higher rates of at least some self-reported health end points (usually GI illness) in relation to water exposure (usually swimmers vs. non-swimmers). That is, recreational water contact by its nature is associated with increased risk of adverse health effects—even if the excess risk is not correlated with increases in fecal indicator organisms.

Both PC and RCT (also called prospective randomized exposure studies [Fleisher et al. 2010]) epidemiology study designs have been used to evaluate recreational waters. The primary difference between RCT and PC design is that in RCT studies the volunteers are randomly assigned to swim in a selected area where the water quality is measured during the timed swimming exposure. A brief overview of each design is provided below.

2.1.1 Prospective Cohort Studies

EPA's current (1986) recreational water quality are based on the observed occurrence of GI illness associated with swimming in fresh or marine recreational waters receiving point sources of effluent from POTWs as determined through several PC studies conducted in the 1970s through the early 1980s. Over the past several years, the EPA conducted a PC study called National Epidemiological and Environmental Assessment of Recreational (NEEAR) Water Study. This series of epidemiology studies evaluated sites at four POTW-impacted Great Lakes beaches and five marine recreational beaches (four POTW-impacted and one non-POTW-impacted). The results of the freshwater and marine studies have been published (e.g., Wade et

al. 2006, 2008, 2010). Further, there have been relatively few studies of inland (non-Great Lakes) waters or of nonpoint source-impacted recreational sites.

2.1.2 Randomized Control Trials

The other major epidemiology study design used for evaluating health risks associated with recreational exposures is the RCT. This design has been used extensively in Europe, and more recently in the United States to determine the association between microbial water quality and increased risk of adverse health effects in swimmers vs. non-swimmers in both marine and fresh recreational waters. The WHO *Guidelines for Safe Recreational Water Environments* (WHO, 2003) apply to both fresh and recreational waters and are largely based on RCT studies conducted in the late 1980s and early 1990s at POTW-impacted marine recreational beaches in the United Kingdom (UK) (Fleisher et al. 1996; Kay et al. 1994). In 2006, the European Union adopted a new directive for the management of bathing water quality that is also based, in large part, on the same UK marine recreational water studies. It also includes the results of a more recent RCT study of POTW- and nonpoint source-impacted fresh recreational waters in Germany (Wiedenmann et al. 2006). Epibathe (2009a, 2009b) describes the results of a series of RCT studies conducted at marine and fresh recreational waters in Europe in 2006 and 2007. Finally, the results of an RCT study of a non-POTW-impacted marine beach in Miami, Florida have been recently published (Fleisher et al. 2010).

2.1.3 Studies that have Generated Health Effects Relationships

A number of epidemiology studies have established a significant relationship between indicator organism density and increased GI illness (Table 2; see also WERF, 2009). Highlighted rows in Table 2 indicate relevant and available epidemiological datasets that were obtained and evaluated in this report (see Section 3). It is beyond the scope of this report to list all epidemiology studies and datasets of recreational waters that have been conducted.

Among the studies listed in Table 2, three have established relationships for inland waters, and all of those studies were of waters that are likely predominantly impacted by human fecal pollution sources. A number of the unique features for the studies that resulted in health effects relations are the following:

- the USEPA (1983, 1984) studies related illness rates to season-averaged culture indicator densities;
- the USEPA (1984) study health effects relationship was developed using pooled data collected for both inland and Great Lakes beaches;
- the USEPA (1983, 1984) studies used different definitions of GI illness than the NEEAR studies;
- the study by Marion et al. (2010) was conducted on a relatively small inland lake, though as shown in Section 3.1, the health effects observed in that study were similar to those observed by USEPA (1984).

In contrast to studies of waters potentially impacted primarily by POTW effluent, epidemiology studies of marine and freshwater non-POTW impacted-recreational waters tend to yield weak associations between increased densities of traditional indicators (*Enterococcus* and *E. coli*) and health risk. Calderon (1991) found no association of indicator density with incidence of adverse

health effects at a non-POTW-impacted pond with no known human fecal pollution effects. Similar results were reported by Colford et al. (2007) for a California coastal beach suspected to be affected primarily by birds, and for a subtropical coastal marine beach most likely affected by dog and human nonpoint sources of fecal pollution (Abdelzaher et al. 2010; Fleisher et al. 2010; Sinigalliano et al. 2010). Results from a study of stormwater impacts on GI illness rates are difficult to interpret (Haile et al. 1999). It is possible that densities of alternative (non-traditional) indicators might be associated with risk and that relationships may be established as a result of ongoing epidemiology studies.

Table 2. Epidemiology studies that have established relationships between FIB density and excess GI illness due to swimming

Study(s)	Indicator and primary detection method	Setting and sources	Study type
Marion et al. (2010)	*E. coli* culture	U.S. impounded freshwater beaches, human and other sources	PC
NEEAR	Density of *Bacteroidales* measured by qPCR	U.S. marine beaches, POTWs	PC
NEEAR (Wade et al. 2006, 2008)	Density of *Enterococcus* spp. measured by qPCR	US freshwater beaches (Great Lakes), POTWs	PC
NEEAR (Wade et al. 2006, 2008)	Density of *Enterococcus* spp. measured by qPCR	U.S. marine beaches, POTW	PC
Wiedenmann et al. (2006)	*Enterococcus* and *E. coli* chromogenic substrate	German, freshwater beaches with one of more point- (including POTWs) and nonpoint sources	RCT
Fleisher et al. (1996) and Kay et al. (1994)	*Enterococcus* culture	U.K. marine beaches, POTWs	RCT
USEPA (1984)	*Enterococcus* and *E. coli* culture	U.S. Great Lakes and inland freshwater beaches, POTWs	PC
USEPA (1983)	*Enterococcus* culture	U.S. marine beaches, POTWs	PC

2.1.4 Factors that Preclude Direct Comparison of Epidemiology Studies

Although epidemiology studies may be able to identify a general association between a given fecal pollution source and indicator organism densities by estimating incidence of disease, a major and ongoing concern is that their results may be limited to describing risk only for beaches similar to those evaluated in the epidemiology studies (e.g., similar fecal sources). The ability to conduct valid comparisons between PC and RCT epidemiology studies depends on several key and potentially interrelated factors. These factors relate to several critical differences in the details of RCT and PC studies, and include the following:

- the manner in which microbial water quality is associated with illnesses;
- the specific exposure that swimmers experience;

- the definition of (GI) illness and the duration of follow-up;
- the age and make up (e.g., tourists vs. locals) of the subject pool;
- the source of fecal indicators and pathogens in the recreational water, their distribution, and temporal-spatial variability of those distributions; and
- the methods used to enumerate the indicators and the corresponding relationships between the indicators and the pathogens, as measured by those methods.

2.2 Differences in Indicator Performance for Indicators as Measured by Culture-Based and qPCR-Based Methods

Cell counts obtained by qPCR are generally greater, often by orders of magnitude, than those provided by MF analyses of the same samples (e.g., as observed by He and Jiang, 2005). This section begins with a review of select studies illustrating the differences between qPCR and MF methods. Common causes of variations in FIB counts via the different methods, as shown in the literature, are listed below by category and subcategory and are discussed subsequently. Results from studies on paired comparisons of different indicator/method combinations are also highlighted throughout the section.

2.2.1 General Method Performance

The performance of selected analytical methods and the fate and transport of cells and genomic material can be influenced by a number of factors, as summarized in Table 3.

Table 3. Summary of factors affecting culture-based and/or qPCR data for fecal indicator bacteria

Influencing factor	Variables	Outcome	Reference(s)
Time of day for sample collection	Exposure to sunlight (UV radiation)	Diurnal variation in FIB density	Haugland et al. (2005)
Variations in method performance	• qPCR amplification efficiency • Limits of detection • Inter-laboratory variability • Choice of target gene(s) • Sample preparation • Concentration factor	Inconsistencies in data generation and interpretation	Haugland et al. (2005) Khan et al. (2007) Kinzelman et al. (2010; preliminary report)
Bacterial viability	Enumeration of live, dead, and/or VNBC FIB	Under-representation of VNBC FIB using MF; over-representation of viable bacteria using qPCR	Nocker and Camper (2006) Nocker et al. (2006) Bae and Wuertz (2009) Varma et al. (2009)

With regards to the sampling approach, the time of day at which samples are collected has been shown to influence the relationship between culture-based and qPCR data (e.g., Haugland et al. 2005). During EPA's NEEAR studies conducted at freshwater sites in the Great Lakes, water samples were tested for *Enterococcus* by qPCR and MF. Samples were collected three times per day (8:00 AM, 11:00 AM, and 3:00 PM), at multiple depths (waist, shin, and knee), and along three transects at all four beaches being studied. A model relating enterococci qPCR and culture data can be developed using results from all the sites and times, or using data that are selected

because they appear to have similar underlying trends. All samples collected from one of these beaches are shown in Figure 3, along with a 45° line showing perfect agreement. Although one may observe a general trend at high FIB density, at densities below ~500 CFU/100 mL, no trend is apparent and there is no obvious choice for the form of a model relating enterococci density derived from qPCR methods to culturable enterococci density. The influence of sample collection time on the relationship between culture and qPCR FIB density is explored in Section 3.2.1.

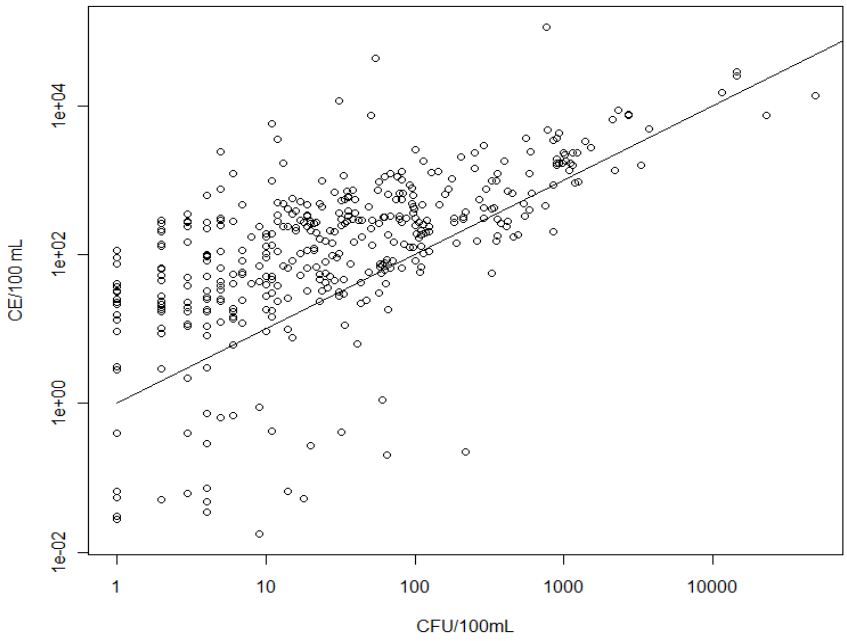

Figure 3. Paired qPCR and culture data for *Enterococcus*: Huntington Beach

Variations in the data because of the performance of the analytical method have been demonstrated in a number of studies. The factors affecting the analytical performance include the efficiency of the qPCR method, the influence of detection limits, and variations in inter-laboratory performance. Haugland et al. (2005) compared a qPCR-based enumeration method for *Enterococcus* to EPA Method 1600 (MF enumeration). The qPCR method employed in that NEEAR-related study demonstrated very high amplification efficiency (0.99, or 99% of amplicons doubling with each cycle) for DNA in dilution water. A modest level of false positive results from qPCR (19% of 217 negative control samples) was speculated to be the result of aerosolized DNA contamination that occurred in analytical laboratories. DNA recoveries in calibration samples (from seeded filters) from the two beaches sampled in the study were 82% and 51% in beach samples, respectively. Results of qPCR and MF enumerations were reasonably well correlated ($R^2 = 0.68$) and log-normal distributions described the densities of *Enterococcus* for samples collected at each beach and for both qPCR and MF enumerations. The authors performed linear regression of qPCR results (as cell equivalents) against MF results (as CFU) for samples collected on multiple days and at multiple locations on two beaches. The resulting relation:

$$\log_{10}(CE) = 1.56 + 0.53\log_{10}(CFU) \quad [1]$$

was proposed to describe the variation in qPCR cell counts with those from MF.

Khan et al. (2007) evaluated qPCR enumeration of *E. coli* in waters of agriculture-dominated watersheds against MF methods, finding as did Haugland et al. (2005), that (1) 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 (2) that standard curves (in this case, based on both dilution water and autoclaved agricultural water) had high coefficients of variation. Khan and colleagues noted that the detection limit in agricultural waters was 10 cells, whereas in dilution water it was 1 cell, and that sample preparation had a profound impact on the viability of the qPCR method. Overall, qPCR enumerations were found to be consistently higher than those from MF methods—the cell counts for all samples analyzed via MF ranged from 1.0 to 2800 CFU/100 mL, whereas the range for qPCR counts was 15 to 9900 cells/100 mL. The authors concluded that qPCR results were less variable than MF results; that qPCR could be used effectively in diverse agricultural watersheds; and that qPCR is more rapid, producing meaningful results far faster than culture-based indicator methods. Application of qPCR methods and interpretation of qPCR results may require consideration of the source of indicator organisms as well as the physiological status of FIB at the time of sampling. For example, the selection of target gene plays an important role in qPCR enumeration. The number of target genes per bacteria cell is dependent upon physiological state of the cell and, for example, during the bacterial logarithmic-growth phase up to 36 rrn genes/cell (range: 12-36 copies/cell) have been reported for *E. coli* (Bremer and Dennis 1996).

Inter-laboratory variability was assessed by Kinzelman et al. (2010; preliminary report) through the analysis of replicate samples in multiple labs to assess the uncertainty and recovery of qPCR. Archived (MF) filters were used to develop replicate samples and qPCR was used to determine *Enterococcus* calibrator cell equivalents (CCEs) in the replicate samples. Multiple laboratories conducted the study and all reported similar findings. A typical plot of the paired samples is presented in Figure 4. In all cases there was very little difference between replicate samples at high indicator density and higher scatter at lower densities. The scatter at low densities appears to be Poisson distributed (analysis not shown) and attributable to sampling uncertainty. The relatively tight distribution of points around the 45° line indicates a consistent and relatively high recovery.

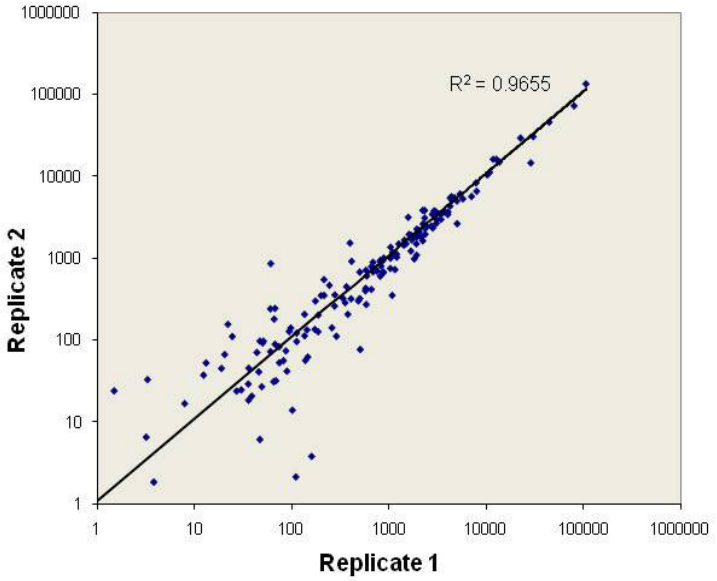

Figure 4. Correlation of qPCR *Enterococcus* densities from replicate samples (Source: Kinzelman et al. 2010)

Comparing some of the key papers used to review general method performance, it is important that the choice of target gene(s) and reaction conditions for the qPCR assays be accounted for when comparing research approaches. As shown in Table 4, there are a variety of bacteria targeted in these key papers, some targeting the genus as a whole (e.g., *Enterococcus* spp.), others targeting a particular species (e.g., *E. coli*), and others targeting a particular strain of a species (e.g., *E. coli* O157:H7).

Accounting for bacterial viability is an important consideration when comparing culture-based methods with qPCR assays. Because of its sensitivity, qPCR will amplify DNA regardless of whether or not it is contained within an intact cell. However, successfully culturing bacteria requires the organism to be intact. There is an intermediate phase termed viable but non-culturable (VBNC), where bacteria appear to not be culturable but retain the ability to reproduce under appropriate conditions. By their nature, culture-based methods will discriminate against free DNA and VNBC bacteria, whereas qPCR methods will quantify genetic material from cells in all states of existence. In the case of detection of DNA from dead cells, this ability is a significant and widely recognized shortcoming of qPCR methods and a potential cause of false positives (NRC 2004).

Several researchers have explored PCR techniques that enable discrimination between DNA from live and dead cells. For example, Nocker and Camper (2006) used a chemical reaction of DNA from dead cells with ethidium monoazide (EMA) to prevent reaction of DNA from dead cells with PCR reagents. Thus, prior to qPCR determination, EMA was added to sample water. EMA, which bonds strongly with DNA and is inactivated in water, can enter only bacterial cells with compromised cell walls. Following this step, DNA was extracted from the live cells using conventional techniques and qPCR was performed for determination of *E. coli* O157:H7 and

Table 4. Summary of the target organisms and primer sets used for a selection of qPCR assays

Reference	Target organism	Target gene	Primers
Haugland et al. (2005)	*Enterococcus*	23R rRNA	ECST748F
			ENC854R
Khan et al. (2007)	*E. coli*	Internal transcribed spacer (ITS) region between the 16S–23S rRNA	IEC-UP
			IEC-DN
Nocker and Camper (2006)	*E. coli* O157:H7	Shiga toxin 1	stx1-forward
			stx1-reverse
	Salmonella enterica	*invA* gene	invA2-F
			invA2-R
Nocker et al. (2006)	*E. coli* O157:H7	Shiga toxin 1	stx1-forward
			stx1-reverse
Bae and Wuertz (2009)	*Bacteroidales*	16S rRNA	BacUni-520f
			BacUni-690r1
			BacUni-690r2
			BacHum-160f
			BacHum-241r
Varma et al. (2009)	*Enterococcus* spp.	16S RNA	ECST748F
			ENC854R
	Bacteroidales	16S rRNA	In Siefring et al. (2008)

Salmonella. The authors suggested that use of EMA with DNA-based methods, if refined, could be a viable alternative to the use of more complicated RNA methods (RNA degrades rapidly after cell death) in developing cell density estimates that exclude dead cells. In a different study, Nocker et al. (2006) suggested that propidium monoazide (PMA) may be a better reagent for use in preventing DNA from dead cells from being amplified in the PCR reaction. In that study, EMA was found to penetrate live cells of some microbiological species, preventing amplification of DNA from the penetrated cells and reducing the accuracy of qPCR estimates of cell density. In contrast, PMA was observed to be selective only for dead cells for the organisms tested in that study.

Similarly, Bae and Wuertz (2009) used PMA to inhibit amplification of DNA from dead cells. The use of PMA (rather than EMA) was suggested because, and as noted by Nocker et al. (2006), EMA is believed to cause degradation of some DNA from viable cells for *E. coli* O157:H7, *Campylobacter jejuni*, *Listeria monocytogenes*, and perhaps other pathogenic and indicator bacteria. Bae and Wuertz (2009) determined PMA reaction conditions (PMA concentration and light exposure time) that optimized removal of target DNA (host-specific *Bacteroidales* genetic markers) from dead cells in wastewater plant influent and effluent. For samples from the wastewater plant effluent, gene copies from qPCR with PMA were only 30%

of those from qPCR without PMA. The difference between qPCR with and without PMA was greater than two orders of magnitude for samples of wastewater plant effluent.

Varma et al. (2009) investigated the difference between PMA-qPCR and qPCR without PMA for *Enterococcus* and *Bacteroidales* in wastewater. Objectives of the study included assessment of water matrix effects on the efficiency of the PMA reaction and exploration of use of qPCR for analysis of wastewater treatment plant operation. Some of the conclusions from that study were the following:

- high levels of biomass or suspended solids in water samples appeared to interfere with the ability of the PMA-qPCR method to specifically detect live cells, and
- standard POTW chlorine disinfection practices resulted in substantially greater reductions in fecal indicator bacteria CFU densities than those observed for PMA-qPCR detectable target sequences.

When reviewing general method performance, the influence of variability in the data generated from different methods and their influence on health effect relations has also been described. Similar to Haugland et al. (2005), albeit for groundwater samples, Lleo et al. (2005) found that qPCR estimates of *E. coli* and *Enterococcus faecalis* density were significantly and consistently higher than those obtained using MF methods. The authors ascribed the difference to the ability of qPCR to detect VBNC cells and suggested that the qPCR method provides estimates of FIB density that are more protective of human health—particularly given the potential for VBNC cells to resuscitate under appropriate conditions.

In the course of EPA's NEEAR studies conducted at Great Lakes beaches, Wade et al. (2008) observed that, in contrast to counts from culture methods, qPCR counts of enterococci were relatively constant during the day. In contrast, counts of enterococci from culture methods show a distinctive diurnal variation, with densities for early morning samples being as much as two orders of magnitude higher than for late afternoon samples. In addition to observing less hourly variation in indicator density measurements using qPCR compared to culturable methods, Wade and colleagues also noted that increasing qPCR densities (expressed as cell equivalents) were associated with excess risk of GI illness in swimmers vs. non-swimmers.

2.2.2 Factors Impacting Indicator/Method Performance

2.2.2.1 Choice of Target Organism

Although this report focuses on the methods used to detect and quantify *Enterococcus* in relation to health effects, it is important to consider how different FIB compare to one another when detected in recreational waters by various methods. Differences in the comparison of qPCR and culture signal exist for different microorganisms, as demonstrated by Rothrock et al. (2009). In that study, triplicate analyses were performed on samples inoculated with dilutions of laboratory cultures of *E. coli* and *C. jejuni*. Plots of qPCR counts vs. MF counts for the two organisms are presented in Figure 5, where the dashed line is a 45° line (exact concordance between methods). In general, the qPCR and culture counts are 1:1, with better agreement at the higher counts (lower dilutions) than at lower counts. The curve for *C. jejuni* has a slope of 1:1, but is shifted above the 45° line. This shift is potentially due to low recoveries typical for this bacterium via culture methods. As suggested in Figure 5, the impact of difference in recovery between the methods is manifested as a shift of the trendline above or below the hypothetical line of perfect

agreement. Differences in recovery among MF counts may results from damage or shock to bacteria during sample preparation, clumping of cells, or difficulty in maintaining growth conditions for some organisms. Differences in recovery among qPCR counts may arise from the presence of substances inhibiting the PCR reaction and differences in specificity, efficiency, and sensitivity for PCR assays targeting different microorganisms.

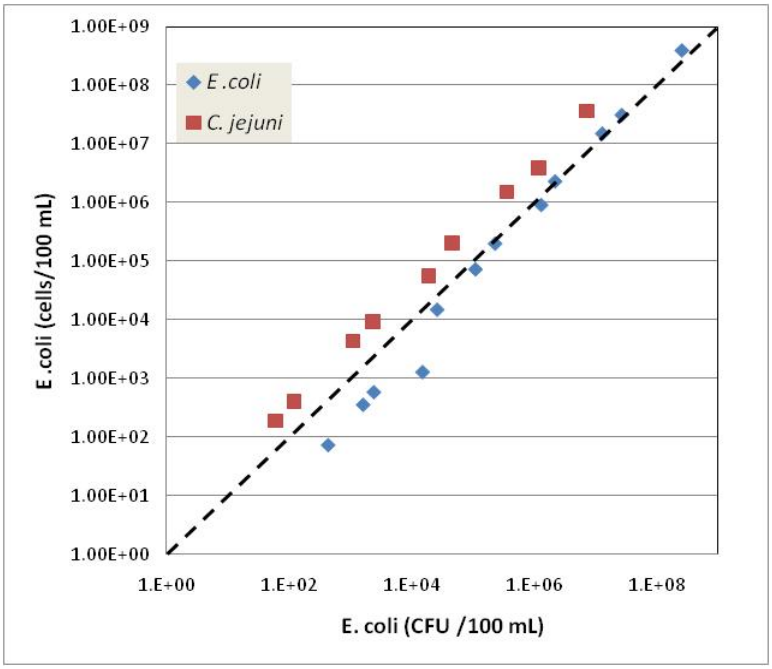

Figure 5. Trend lines for dilutions of laboratory cultures of *E. coli* and *C. jejuni* enumerated using qPCR and culture methods (SOURCE: Rothrock et al. 2009)

Simultaneous detection of *Enterococcus* and *Bacteroidales* in recreational marine water was undertaken by Elmir et al. (2009) using MF and qPCR detection methods. Two pools, one large and one small, were filled with local offshore marine water and recreational users (adults and toddlers) were required to be in the water for pre-determined periods. Using three analytical methods, *Enterococcus* and *Bacteroidales* were enumerated from the water samples before and after users were in contact with the water. Microbial concentrations in the source water were generally low (MF = 5 [std. dev. ± 7] CFU, qPCR = 29 [± 49] genomic equivalent units). *Bacteroidales* was detected only using qPCR human markers (UCD and HF8), yielding genomic equivalent units (GEU) of 45 (± 18 3) and 3 (± 10) GEU/100 mL, respectively. The concentrations of *Enterococcus* and *Bacteroidales* in both pools were variable following recreational use. *Enterococcus* densities calculated using MF ranged from 1.8×10^4 to 2.0×10^6, whereas values calculated using qPCR ranged from 3.8×10^5 to 5.5×10^6. *Bacteroidales* values ranged from below the limit of detection (1.4×10^3 GEU) to 1.3×10^6. The authors concluded that the bathers appeared to release significant amounts of FIB via shedding from their bodies and into the water column. For this study, an added significance of the study is that culture and qPCR counts for enterococci from bathers were generally in the same range. By comparison, effluent from POTWs employing chlorination may have culture and qPCR counts that differ by many orders of magnitude (e.g., He and Jiang 2005, Bolster et al. 2005) and qPCR and culture

counts in POTW-impacted coastal waters may differ by smaller margins than observed in chlorinate POTW effluent (Byappanahalli et al. 2010).

A suite of FIB from brackish water was analyzed by Ortega et al. (2009). Eighteen sampling events were conducted over five sampling trips in the St. Lucie River Estuary, South Florida. Enterococci were enumerated using MF, whereas *E. coli* were assayed using the most probable number (MPN) method. The values for enterococci and *E. coli* were similar (10^1 to 10^2 CFU or MPN/100 mL) and an R^2 value of 0.53 was reported when correlating the two fecal indicators.

Finally, Agudel et al. (2010) used MF and multiplex real time PCR to compare the concentrations of *Bacteroides* spp., *E. coli*, and enterococci in a variety of environmental matrices in Barcelona, Spain. Seventy-four samples from rivers, wells, urban groundwater, and wastewater were analyzed for FIB counts using standardized MF methods for *E. coli*, total coliforms and enterococci, and qPCR assays for total *Bacteroides* spp. and enterococci. Based on the authors' statistical analyses, Figure 6 summarizes the enumeration of the fecal indicators. The authors concluded that bacterial quantification data were more homogeneous when using PCR than when using conventional culture microbiology.

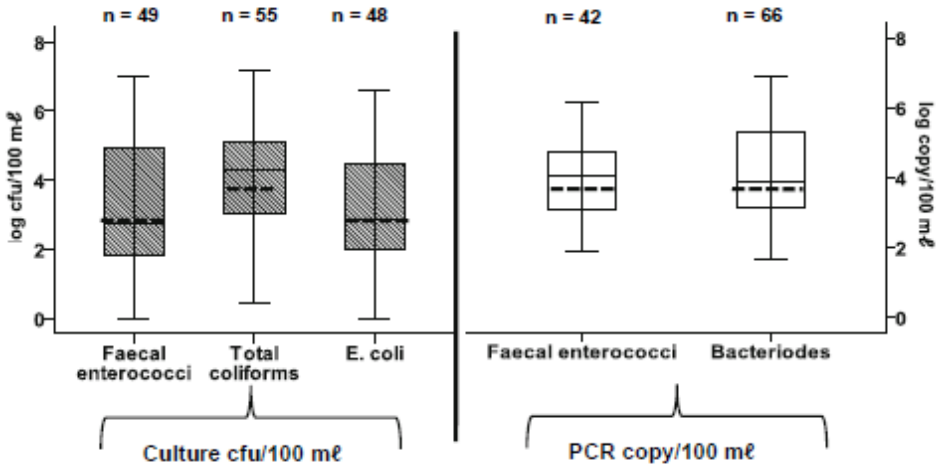

Figure 6. Comparison of bacterial data from culture-based and multiplex real-time PCR methods, and statistical sample description (SOURCE: Agudel et al. 2010)

2.2.2.2 Water Matrix

Inhibition of the qPCR amplification reaction is an important factor to consider when comparing performance of qPCR with the *Enterococcus* MF technique. Duprey et al. (1997) hypothesized possible seawater inhibition of the qPCR assay through secretion of substances by phytoplankton and zooplankton, while Haugland et al. (2005) also noted frequent inhibition of qPCR in undiluted Great Lakes water. However, such inhibition is reduced or eliminated by performing serial dilutions of the sample (Ahmed et al. 2009).

Characteristics of the water matrix also impact quantification assays. Sinton et al. (2002) showed that salinity increased bacterial cells rate of decay, which would widen the already demonstrated

gap between culture-based and qPCR results in marine waters. Duprey et al. (1997) noted in their seawater study that DNA persistence was lower during the summer for free and dead cell DNA, and could extend up to 55 days in the winter for dead cell DNA. The persistences of naked DNA and a small inoculum of dead cell DNA were the lowest at two days. Walters et al. (2009) reported that in sewage microcosms exposed to sunlight, culturable *Enterococcus* concentration fell below detection limit within 5 days while qPCR signal persisted for at least 28 days. These observations are consistent with the general observation of higher qPCR-based bacterial densities (enumerated in cell equivalents) than culture-based counterparts in both marine and freshwaters (Byappanahalli et al. 2010; Bower et al. 2005; Haugland et al. 2005; Lavender and Kinzelman 2009; Morisson et al. 2008; Whitman et al. 2010)—except for Elmir et al. (2009) who found comparable densities from both techniques in a marine setting and Noble et al. (2010) who found underestimation of bacterial density via qPCR for marine settings.

The two studies by Noble et al. (2010) were interpreted by the authors as demonstrating an underestimation of *Enterococcus* by qPCR with regards to culture-based methods in marine water. However, on this point there are several concerns in interpretation that need to be considered. The primer set used was not specifically described in the study. In the discussion it was indicated that the primers used were designed for high specificity for *E. faecalis* and *E. faecium* (which would detect a narrower population of *Enterococcus*) in this study than in studies that used the EPA method. The use of a single species calibrator standard could also result in disagreement between the techniques, as different species of enterococci may have different numbers of the target gene. Finally, the authors pointed that the relative quantification method with a calibrator and a salmon testes DNA control used by other teams expressing results in CCEs could explain the observed discrepancy between underestimation and overestimation. They discarded the possibility of error originating from PCR chemistry by showing no significant difference between the Taqman and Scorpion assays for *Enterococcus*. However, this team confirmed the previously published findings of significant correlation between qPCR and culture-based assays for both *Enterococcus* and *E. coli*, with a stronger agreement for *E.coli* than for *Enterococcus* in terms of beach management decisions (88% vs. 94%). With regards to technique implementation, the authors showed that microbiologists with little qPCR training produced similar results as their experienced counterparts, while finding DNA extraction complex and time-consuming. This surprising finding was attributed to a simplified PCR preparation protocol.

The forgoing findings illustrate the complexities of applying qPCR methods, particularly since this technique can amplify all DNA present in a sample, regardless of its viability state. This implies that ambient background DNA (i.e., naked and dead cell DNA; see Lavender and Kinzelman 2009) needs to be assessed when employing qPCR methods, regardless of the water matrix. Lavender and Kinzelman (2009) successfully demonstrated the use of a site-specific corrective factor to correlate culture and qPCR counts and highlighted its importance for the prediction of beach microbial water quality status.

2.2.2.3 Primary Fecal Pollution Source on Indicator/Method Performance

Fecal pollution sources can include point- and nonpoint sources. Common point sources are POTW effluents and stormwater outfalls, while nonpoint sources include wildlife, human shedding and soils and sediments. Although most U.S. POTW effluents are disinfected, some are not, and in some facilities seasonal disinfection is employed. Disinfection usually yields

minimal to null culture-based results for FIB, while qPCR results can still remain high, accounting for the VBNC and background fractions of enterococci and other FIB. However, POTW monitoring results typically represent human-derived enterococci, with direct relevance to public health. Comparatively, indicators in stormwater outfalls can originate from a variety of sources, in which case qPCR primers are chosen based on an understanding of the origin of the fecal pollution. Such human/nonhuman speciation would be especially necessary in rural settings with agriculturally developed watersheds but could be recommended at all recreational sites due to potential contribution of native *Enterococcus* to beach advisories (Yamahara et al. 2009). Simultaneous probing with *Enterococcus* and *Bacteriodales* primer sets can provide such speciation (Converse et al. 2009; Elmir et al. 2009; Shanks et al. 2009). The human shedding of enterococci from bathers was considered minimal in marine waters during initial load or bathing cycle (about 4%) and proportional to the body's surface area (Elmir et al. 2009). These researchers noted that additional bathing cycles could increase the predominance of human shedding of FIB.

2.2.2.4 Treatment and Environmental Factors

Lavender and Kinzelman (2009) and Varma et al. (2009) investigated the effects of wastewater treatment on culture-based and qPCR results for enterococci. Their similar findings showed that culture-based numbers are strongly reduced (2 to 5 orders of magnitude), especially by secondary treatment and disinfection, while qPCR numbers experience smaller reductions or remain unchanged. As noted previously, this is probably due to the strong impact of disinfection on culturable cells. Environmental factors (e.g., rain events, UV light) can also significantly affect qPCR and MF results differently. Lavender and Kinzelman showed that culture-based *Enterococcus* densities at stormwater outfalls were significantly different depending on rain events and location, while no significant differences were observed for qPCR. These researchers also showed that ambient background DNA appeared insignificant during stormwater events. Varma et al. (2009) also found that qPCR results were virtually unaffected between normal and stormflow operations. The contribution of FIB from undeveloped watersheds to reference beaches studied by Griffith et al. (2010) implies that native sands are at least partially responsible for some FIB exceedances of water quality thresholds. Furthermore, Walters et al. (2009) showed that sunlight seems to inhibit the cultivability of cells and accelerate their degradation, reducing persistence of enterococci by half for culture-based results. Neither dead cell nor naked DNA (ambient background DNA) appeared to be affected by sunlight. Given the nature of qPCR and the relative small fraction of viable and culturable FIB, qPCR are not greatly affected by insolation while culture-based methods experience strong reduction by photoinactivation during sunlight exposure.

2.2.3 Uncertainty of Indicator/Method Combinations at Low Density

As noted previously, qPCR measurements of FIB exhibit higher variability at low densities (Haugland et al. 2005). This trend is probably explained by the presence of ambient background DNA, the abundance of which is site-specific. Lavender and Kinzelman (2009) proposed to reconcile qPCR counts with culture counts by introducing a site-specific corrective factor in the calculation of qPCR-derived densities. Bae and Wuertz (2009) proposed a more universal solution by modifying the traditional qPCR approach with the introduction of PMA, allowing qPCR to distinguish viable from non-viable cells. This approach was found to be promising and was used successfully by Varma et al. (2009).

3 Analyses

This section describes and demonstrates analyses comparing the illness response to exposure relationship for multiple indicator/method combinations used in epidemiology studies and evaluating multiple indicator/method combinations to develop quantifiable relationships. The approach taken to meet the report's objectives is through the use of available data to provide a proof of concept for the application of various statistical techniques to develop equivalence between indicator densities for different indicator-method combinations. These initial analyses illustrate the most direct approaches for linking indicator-method combinations and the steps that can be undertaken to effect meaningful comparison of results across disparate datasets. Two approaches for developing quantifiable relationships for "linking" indicator/method combinations, the Risk Link approach and Water Quality Link approach, are demonstrated in this section.

The Risk Link approach is evaluated first using three different statistical linkage analyses. The first is a straight-forward linkage of *Enterococcus* density as measured by qPCR to *Bacteroidales* density as measured by qPCR. The second demonstration shows analyses that can be conducted prior to linkage of indicators using curves from different studies. That demonstration links curves (Marion et al. 2010 and USEPA 1984) based on the same indicator-method combination (*E. coli* by MF), but from epidemiological studies conducted in different settings, with different samples sizes, and at different times. The third demonstration illustrates the linkage of *Enterococcus* enumerated by MF to *Enterococcus* enumerated by qPCR, both from studies conduted in POTW-impacted marine waters. That demonstration also illustrates the translation of GI definitions between studies, prior to conducting the Risk Link analysis. Comparisons beyond the three Risk Links demonstrated here are possible, and additional linkages may be conducted as alternative statistical analyses or datasets become available.

Secondly, the Water Quality Linkage approach is evaluated using two statistical models, simple linear regression and broken stick (segmented) regression. Using paired water quality data from the NEEAR freshwater dataset, the culture-qPCR relationships appear to vary significantly among beaches. The linkage results in translated culture *Enterococcus* criteria values generally significantly higher than the existing culture criteria for those specific beaches. Other site-specific models, or models that incorporate data beyond paired indicator-method data, may be developed to overcome the limitations of the evaluated regression models in the future.

3.1 Risk Link

Several Risk Link demonstrations are provided in this section. The simplest and perhaps strongest application of the Risk Link approach is the linkage of densities of indicators for multiple indicator-method combinations via health effects curves generated in the same epidemiology study. In this comparison, illness definitions, sampling strategies, statistical interpretations and fecal pollution sources are the same, and no conversions or intermediate data analyses are necessary. However, Risk Linkages can also be made using health effects curves from different epidemiology studies. In that event, analyses must be conducted prior to the linkage of indicator densities, resulting in linkages that are less direct than those established using multiple health effects curves from a single epidemiology study.

The Risk Linkages demonstrated in this section illustrate both the direct application of the Risk Link using health effects relationships from the same study, and the less direct Risk Link, that requires preliminarily conversions or analyses. The first demonstration illustrates the linkage of *Bacteroidales* as measured by qPCR and *Enterococcus* as measured by qPCR. Health effects curves for both of those indicator-method combinations were developed in NEEAR study marine beach epidemiology studies in which both indicator densities were measured concurrently and associated with the same illness levels. The second demonstration illustrates techniques for assessing the differences in health effects associations with indicator densities for a common indicator-method combination (*E. coli* by MF) for epidemiology studies with similar designs, but conducted at different times and in different settings. The two studies from which health effects curves were drawn are a recent study of illness-indicator density associations at a small inland lake (Marion et al. 2010), and the studies conducted as a prelude to establishment of the current RWQC (USEPA 1984). The third demonstration illustrates the harmonization of health effects curves, from two epidemiology studies using different illness definitions, prior to the application of the Risk Link approach. In that study, qPCR-based *Enterococcus* health effects curves from the NEEAR study marine beaches are linked to health effects curves based on *Enterococcus* as measured by MF in the USEPA (1983) epidemiology studies. Implicit in the third demonstration is an assumption that the conditions at the beaches studied in the two epidemiology studies were sufficiently similar to allow meaningful comparison of their health effects curves.

3.1.1 Review of NEEAR Study and Other Health Effects Relationships for qPCR and Culturable FIB

The NEEAR studies provide the best current association between water quality measures and illness rates (i.e., health effects curves), as they are the most recent, largest, and most carefully designed epidemiology studies conducted by EPA. In this section, health effects relations from the NEEAR studies are presented. Those relationships may be used in development of new qPCR-based criteria. Note that the results presented in this report are only a small portion of the findings of the NEEAR studies. More detailed results and full descriptions of the studies can be found in Wade et al. (2006, 2008, 2010).

To date, EPA's NEEAR recreational water epidemiology studies have employed both qPCR and culture-based methods in determining FIB density. Further, the studies have yielded health effects relationships between *Enterococcus* density as measured by qPCR, *Bacteroidales* density as measured by qPCR, and health effects (GI illness) for both marine and Great Lakes POTW-impacted sites. The development of associations of illnesses with those indicator-method combinations is particularly significant because qPCR is a rapid method. Health effects curves generated from observed indicator densities in the NEEAR study marine beaches and for the USEPA (1983) study are presented in Table 5. Candidate NEEAR study health effects relations for marine beaches were provided in a personal communication from T. Wade (2010).

Table 5. Summary of health effects relationships from USEPA NEEAR studies and marine studies conducted to support development of the current RWQC

Setting	Health effects relationship	Indicator and method	Reference
Great Lakes	$N/1000 = 0.0214 \log_{10}(C_{ENT}) - 0.00918$	*Enterococcus* measured by qPCR (CE/100 mL)	Wade et al. (2006)
Marine beaches	$N/1000 = 0.0517 \log_{10}(C_{ENT}) - 0.101$	*Enterococcus* measured by qPCR (CCE/100 mL)	T. Wade, personal communication, unpublished data, 2010
Marine beaches	$N/1000 = 0.00406 \log_{10}(C_{ENT}) + 0.0144$	*Enterococcus* measured by MF (CFU/100 mL)	T. Wade, personal communication, unpublished data, 2010
Marine beaches	$N/1000 = 0.0368 \log_{10}(C_{BAC}) - 0.0962$	*Bacteroidales* measured by qPCR (CCE/100 mL)	T. Wade, personal communication, unpublished data, 2010
Marine beaches	$N/1000 = 0.0219 \log_{10}(C_{ENT}) - 0.01485$	*Enterococcus* measured by MF (CFU/100 mL)	USEPA (1983)

In the EPA study of marine beaches impacted by POTW effluent, Wade et al. (T. Wade, personal communication, unpublished data, 2010) found that swimmers experienced more GI illness than non-swimmers on days when *Enterococcus* density (as measured by MF) exceeded the current geometric mean guidelines. However, the association among swimmers was not statistically significant. The authors contrast this lack of association with those for *Enterococcus* and *Bacteroidales* enumerated by qPCR, both of which were associated with a consistent trend of excess GI illness among swimmers. Note that a weak association of *Enterococcus* density as measured by MF with health effects (GI illness) was also observed in EPA's Great Lakes epidemiology studies (Wade et al. 2006, 2008).

For a rate of swimming-associated GI illness of 35 per 1000 swimmers (risk levels similar to those underlying the USEPA1986 criteria) the Great Lakes health effects curve indicates an indicator density of 116 qPCR *Enterococcus* CCE/100 mL. For a swimming-associated GI illness rate of 35 per 1000 swimmers, the marine health effects curve based on qPCR enumeration of enterococci yields a corresponding indicator density of 427 qPCR *Enterococcus* CCE/100 mL. The culture-based NEEAR study health effects curve for marine waters produces a high FIB density corresponding to a swimming-associated illness rate of 35 per 1000 swimmers.

3.1.2 Direct Application of the Risk Link Approach: Linking *Bacteroidales* Measured by qPCR to *Enterococcus* Measured by qPCR

Application of the Risk Link approach to determine densities of *Enterococcus* measured by qPCR and *Bacteroidales* measured by qPCR that are consistent with comparable risks is straightforward. The *Enterococcus* and *Bacteroidales* health effects curves presented in Table 5 were generated using water quality and attributable illness rate data collected concurrently at the same beaches and from the same cohorts. Thus the health effects curves (Table 5) may be used without prior analyses to establish risk-linked indicator densities. Risk-linked *Enterococcus* and *Bacteroidales* densities for marine waters at three risk levels of interest are presented in Table 6

Table 6. Risk-linked *Bacteroidales* and *Enterococcus* densities

Tolerable attributable illness level (as NGI[†] per 1000 swimmers)	*Enterococcus* density, qPCR enumeration (CCE/100 mL)	*Bacteroidales* density, qPCR enumeration (CCE/100 mL)
35	427	3680
43	610	6060
82	3460	69,600

[†] NEEAR study definition of GI illness

3.1.3 Linking *E. coli* Densities as Measured by MF via Health Effects Curves from Different Epidemiology Studies

In the second application of the Risk Link approach, health effects curves based on densities of *E. coli* measured by MF are used to demonstrate the viability of comparing results from studies conducted at different times and in different settings. The analyses performed in this demonstration are intended to illustrate techniques for assessing whether epidemiology study findings differ between studies and provide insights regarding the combinability of data from multiple studies. The demonstration relates to the P15 objectives in that it illustrates methods for comparing the GI illness response to exposure relationships, though in this case the relationships are for the same indicator-method combination. Note that the analyses presented are not the only ones that could be performed. Other meta-analyses that have explored the combination of health effects data from disparate epidemiology studies are summarized in a text box concluding this section.

A recent epidemiology study conducted by Marion et al. (2010) resulted in a health effects curve relating *E. coli* density via MF to both HCGI and GI illness. Limitations of this study's use in developing general, quantifiable relations are that the study was relatively small (minimum number of swimmers on a study day was 11 and the maximum was 88); the study was conducted for an inland waterbody (impounded reservoir); the dominant fecal pollution source was not clearly POTW discharge; and *E. coli* by MF was the sole indicator/method investigated. Despite these limitations, the analyses seek to evaluate whether the health effects curves generated in the Marion study are comparable to those generated in the epidemiology studies conducted by USEPA (1984) on freshwaters in support of the 1986 criteria. Although not conclusive, the comparison indicates that for the beaches studied by USEPA (1984) and Marion et al. (2010), FIB densities are associated with similar risks.

The Marion et al. (2010) and USEPA (1984) studies differ in the following respects:

- the decade in which the study was conducted,
- the manner in which water quality and health effects are characterized, and
- the settings at which data were collected.

Plots of health effects vs. indicator densities for the two studies are presented in Figure 7. Data from the USEPA studies are blue diamonds and the regression line through the data is blue. The data from Marion and colleagues are presented in two different ways. In that study, illness rates are provided for ranges of FIB densities. Those data are plotted as purple crosses. Data for each study day were provided by the author in a personal correspondence (Marion, personal communication, 2010) and those data and associated regression line are shown as red circles and a red line. Individual study day data from the Marion et al. (2010) study show wide scatter, which is not surprising given the overall low rates of attributable illness and the relatively small number of swimmers on each study day. Overall, the ranges of attributable illness rates for the two sets of studies are similar (based on the binned data from the Marion study) and the slopes of the health effects curves also appear similar.

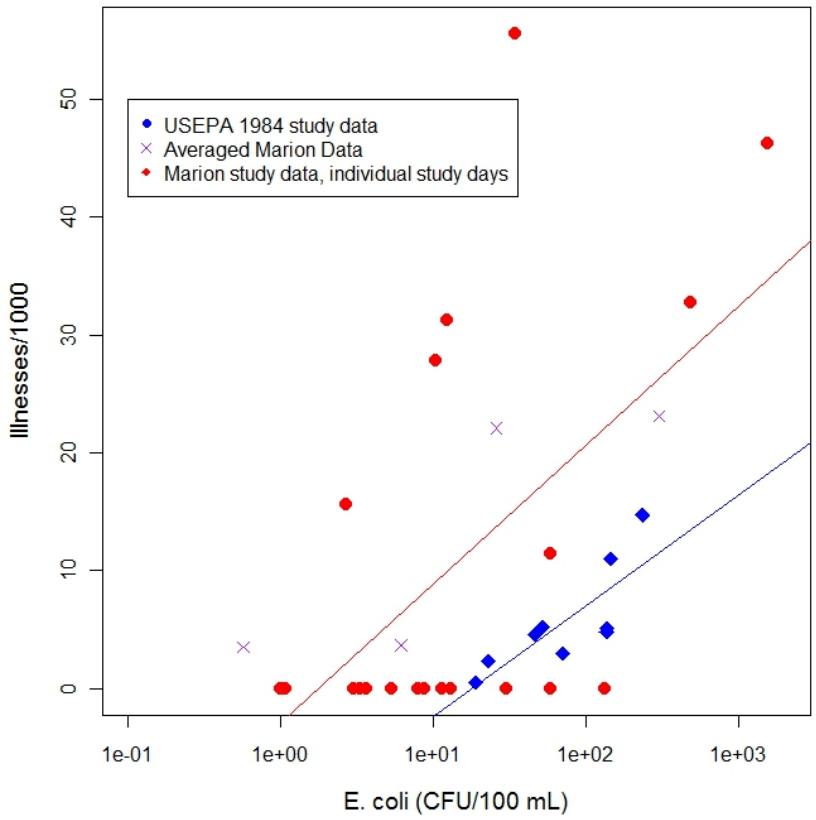

Figure 7. Health Effects Data and Trend Lines from the USEPA and Marion Studies

To test whether the data from the two studies indicated the same or different trends, an analysis of covariance (ANCOVA) for the individual study day and USEPA datasets was performed. The ANCOVA tests the hypothesis that the slope and intercept estimates for regression lines are the

same, accounting for covariation of the slopes and intercepts. The results of the ANCOVA are presented in Table 7. In ANCOVA analyses, the p-value is the probability that the parameter is the same for both models and a p-value above a value of 0.05 indicates that the null hypothesis is accepted. For the two epidemiological datasets the p-values for both the slopes and intercept are significantly above 0.05. Thus, it is concluded that the health effects curves are not different from each other and that the data may be pooled.

Table 7. Results from ANCOVA comparing data from the USEPA (1984) and Marion et al. (2010) epidemiology studies

Parameter	P-value for comparison between models
Slope	0.85
Intercept	0.10

The finding that the health effects models for the USEPA and Marion epidemiology studies are not significantly different has two ramifications. First, the study by Marion et al. (2010) was conducted at an inland lake (impounded reservoir) with mixed fecal pollution sources, one of which was discharge from small POTWs into tributaries to the lake. The studies conducted by USEPA (1984) took place on Lake Erie and at a freshwater inland lake (Keystone Lake, Oklahoma). While not conclusive, the ability to pool health effects data from these sites indicates that health effects at POTW-impacted inland waters are similar to those for POTW-impacted coastal waters. Second, the ANCOVA supports health effects relations for the two studies being similar despite the long time period between the studies. This finding cannot, however, be used to assert that water quality at all sites has not changed in the intervening years between the studies. However, the similarity in the health effects curve for the Marion et al. (2010) and USEPA (1984) studies may indicate that the earlier relationship is indicative of indicator-health effects relationships that may be observed for some types of current recreational sites.

Again, it is noted that the comparison of the USEPA (1984) and Marion et al. (2010) studies is conducted to demonstrate techniques for establishing Risk Links. In this case, the analyses are conducted to establish that indicator densities are associated with similar illness levels in different studies. Analyses such as these may be required to establish linkages in the event straightforward linkages with data from a single study or set of studies are not possible.

Finally, we note that meta-analyses of epidemiology studies have evaluated either implicitly or explicitly the similarities and differences in observed health effects and in health effects associations with indicator densities for different beaches and times. The objectives of those meta-analyses are in some respects similar to the analyses conducted to establish that the USEPA (1983) health effects curves are generally applicable today. Brief descriptions of other meta-analyses that have been performed are provided in the text box below.

Prüss (1998) conducted a systematic review following discussions between the WHO Regional Office for Europe and WHO Headquarters to initiate development of new guidelines for recreational use of the water environment. The comprehensive review of 22 published studies on sewage pollution of recreational water and health outcomes concluded that there was a causal association between GI illness symptoms and increased bacterial indicator density (i.e., enterococci for marine, enterococci and *E. coli* for fresh) in recreational waters.

A meta-analysis of 18 published studies (Zmirou et al. 2003) provided a scientific basis for establishing new standards for the microbial quality of marine and fresh recreational waters to replace the 30 year-old European Union bathing water quality guidelines. The researchers provided four major results: (1) increased concentrations of fecal coliforms or *E. coli* and enterococci in both fresh and marine recreational waters are associated with increased risks of acute GI illness, with enterococci eliciting four to eight times greater excess risks than fecal coliforms or *E. coli* at the same concentrations; (2) GI illness risks associated with enterococci occur at lower concentration in marine versus fresh recreational waters; (3) increased concentrations of total coliforms have little or no association with GI illness risk; and (4) no evidence exists of a threshold of indicator density below which there would be no GI illness risk to bathers.

Wade et al. (2003) conducted a systematic review and meta-analysis of 27 published studies to evaluate the evidence linking specific microbial indicators of recreational water quality to specific health outcomes under non-outbreak (endemic) conditions. Secondary goals included identifying and describing critical study design issues and evaluating the potential for health effects at or below the current regulatory criteria (USEPA, 1986). The researchers concluded that (1) enterococci and to a lesser extent *E. coli* are adequate indicators (predictors) of GI illness in marine recreational waters, but fecal coliforms are not; (2) the risk of GI illness is considerably lower in studies with enterococci and *E. coli* densities below those established by EPA (1986), thus providing support for their regulatory use; (3) *E. coli* is a more reliable and consistent predictor of GI illness than enterococci or other indicators in fresh recreational waters; and (4) based on heterogeneity analyses, studies that used a non-swimming control group and that focused on children found elevated GI illness risks.

3.1.4 Application of the Risk Link Approach for Health Effects Curves Based on Different Illness Definitions

Two steps are required to demonstrate the Risk Linkage between *Enterococcus* measured by MF in studies of POTW-impacted marine sites conducted by USEPA (1983) and *Enterococcus* measured by qPCR in NEEAR marine water studies:

- illness rates from the USEPA (1983) studies are translated to equivalent rates reflecting different illness definition used in NEEAR studies, and
- demonstration that the association of indicator densities and health effects observed in the two studies are similar.

Direct comparison of health risks associated with water quality published in 1986 and the health, water quality relationships developed by Wade (2009 to 2010) cannot be done because the case definitions underlying the relationships are not the same. Wymer (L. Wymer, personal communication, 2010) suggested a translation algorithm for converting the health criteria developed in 1986, based on a HCGI case definition, to an equivalent health criteria that uses NEEAR study data based on a NGI case definition. The translation would be independent of water quality data and uses health data from the 1972 to 1981 and 2002 to 2009 EPA studies.

Conversion of the 1986 criteria to potential new criteria requires only the illness rates (using their respective case definitions) from the non-swimming populations from both study periods

and the 1986 criteria value. The algorithm uses the HCGI illness rate (*IR*) in non-swimmers and the acceptable HCGI illness rate from the criterion, and calculates a relative risk (*RR*) value as:

$$RR_{1986} = \frac{\text{Criteria acceptable } IR_{1986} + \text{Non - swimmer } IR_{1986}}{\text{Non - swimmer } IR_{1986}} \quad [2]$$

An equivalent criterion value (ECV) can be calculated by multiplying the relative risk from the 1986 data by the 2002-2009 non-swimmer illness rate (NGI) and subtracting the 2002-2009 non-swimmer illness rate (NGI) from that value.

$$ECV = \left(RR_{1986} \times \text{Non - swimmer } IR_{2009}\right) - \text{Non - swimmer } IR_{2009} \quad [3]$$

Non-swimmer illness rates observed in the epidemiology studies conducted to support the1986 criteria and in the NEEAR studies were 13.6/1000 swimmers and 59/1000 swimmers, respectively (USEPA 1986). Based on the translation and observed non-swimmer illness rates, equivalent criteria values for three risk levels of interest are provided in Table 8.

Table 8. Equivalent Criteria in Terms of HCGI and NGI Definitions of Illness

HCGI criterion (per 1000 swimmers)	ECV (NGI equivalent) (per 1000 swimmers)
8	35
10	43
19	82

At present, EPA has not selected new RWQC based on qPCR or MF measurement of *Enterococcus*. Therefore, this section concludes with illustration of the risk linkage approach for a range of risk levels and approaches for selecting criteria. Reasonable levels of risk for EPA to evaluate when selecting new criteria are the tolerable illness rates from the 1986 RWQC for recreation in freshwaters (8 HCGI attributable illnesses per 1000 swimmers), the tolerable illness rate for recreation in marine waters (19 HCGI attributable illnesses per 1000 swimmers), or some intermediate risk level (e.g., 10 HCGI attributable illnesses per 1000 swimmers).

The FIB densities used in the USEPA (1983) health effects relationship are geometric mean values for indicator densities taken over entire recreation seasons. Recognizing that water quality varies over the recreation season, current criteria for single sample maximums for designated beaches are based on the 75th percentile of a log-normal distribution whose geometric mean corresponds to the indicator density from the USEPA health effects relationship and whose log-transformed standard deviation is 0.7. New or revised criteria for *Enterococcus* using MF methods can be based on the 75th percentile value or other values selected based on improved characterization of the temporal variability of *Enterococcus* densities in typical marine waters.

Considering all of these and other factors, example criteria values for qPCR and culture-based datacorresponding to equivalent risk levels are presented in Table 9. The qPCR example criteria were calculated using the NEEAR study health effects relationship for all marine POTW-

impacted beaches combined. Note that all alternatives to the current culture-based criteria value for designated beaches (104 CFU/100 mL) are significantly lower than the current RWQC.

Table 9. Examples of equivalent criteria values for marine waters via the statistical linkages approach

Tolerable attributable illness level (as HCGI per 1000 swimmers)	Tolerable attributable illness Level (as NGI per 1000 swimmers)	*Enterococcus* density measured by qPCR (CCE/100 mL)	Geometric mean *Enterococcus* density measured by MF (CFU/100 mL)	75th percentile value† (CFU/100 mL)
8	35	427	11	33
10	43	610	14	42
19	82	3460	35	104

† The 75th percentile value for *Enterococcus* density based on the calculated geometric mean density and assuming a typical standard deviation of log-transformed density for marine sites of 0.7.

3.2 Water Quality Link Approach

The Water Quality Link approach requires two steps: (1) development of a model relating indicator density as measured by one indicator-method combination to indicator density as measured by another and (2) establishment of equivalent values of indicator densities based on two indicator-method combinations using the resulting model and a health effects curve for one of the indicator-method combinations. As noted in Section 2, approaches for establishing models relating indicator-method pairs are in development. Recently published studies have included models based on simple linear regression (e.g., Noble et al. 2010; Whitman et al. 2010) and suggested correction factors that account for site- or time-specific conditions influencing the ratio of indicator counts by different methods (e.g., Lavender and Kinzelman 2009).

This section demonstrates both the modeling component and the linkage component of the Water Quality Link approach. First, general features of the paired data used in the Water Quality Linkage are presented. Next, development of the statistical models (functional form of the relationship between the densities by the different methods and data used to determine the model parameters) for linking indicator-method density pairs is demonstrated. The demonstration includes techniques for determining which data to use in generating the model and alternative model forms for NEEAR study freshwater beach data pairs for *Enterococcus* density as measured by qPCR and *Enterococcus* density as measured by MF. ANCOVA (described in Section 3.1.3) is suggested as a potential means for screening data prior to model development. Simple linear regression and broken stick (segmented) regression models are fit to the data and compared. The broken stick regression models were investigated because simple linear regression models of the NEEAR study data exhibited distributions of residuals that indicate the simple linear models may not adequately describe the relationship between the indicator method combinations.

3.2.1 Datasets Used in Water Quality Link Demonstrations

The NEEAR study freshwater data are used in demonstrating the Water Quality Link approach. Briefly, the NEEAR study water quality data include paired data for *Enterococcus* as measured

by MF, and *Enterococcus* as measured by qPCR for samples collected at multiple transects at four beaches and at three sample collection times for each sample day. The four beaches included Huntington Beach (Lake Erie), Washington Park Beach (Lake Michigan), Silver Beach (Lake Michigan) and West Beach (Lake Michigan). Samples were collected along multiple transects on each sample day at 8:00 AM, 11:00 AM, and 3:00 PM and at shin and waist depths. Other water quality and environmental data were collected during the NEEAR studies, but those data were not used in the current demonstration.

3.2.2 Linear Models of Log-Transformed Indicator Data

The most critical component of the Water Quality Linkage approach is establishing a model relating the densities of paired water quality samples for different indicator/method combinations. Several researchers have proposed linear models for relating log-transformed *Enterococcus* qPCR densities and culture counts (e.g., Haugland et al. 2005; Noble et al. 2010; Whitman et al. 2010;). This section describes a process by which a linear model of log-transformed data may be developed. Model development includes regression of data to find model parameters (slope and intercept). Additionally, the data undergo a selection process by which similar data are included, and dissimilar data are excluded, in the analysis.

To introduce the Water Quality Link approach and modeling of paired indicator density data, a plot of paired data (*Enterococcus* density as measured by MF and *Enterococcus* density as measured by qPCR, Figure 8) is presented and described. The data and fits shown in this example provide background for development of the linear regression model and orient the reader to general features of the relationship between densities of *Enterococcus* measured by MF and *Enterococcus* measured by qPCR. The data shown in Figure 8 were collected and analyzed by Ferretti et al. (2008) in a study of multiple beaches along the New Jersey shore. They show a general trend toward a 1:1 correspondence between cell equivalents (CE) and CFUs at high indicator densities, with significantly greater scatter, and possibly a different trend, at low indicator densities. These trends are both expected and, as described in Section 2, are likely due to the following:

- the association of each live, culturable cell with genetic material measured via qPCR;
- high uncertainty in qPCR at low densities due to analysis of very small samples volumes (in terms of original sample volume);
- divergence in the culture and qPCR signals as organisms age; and
- the difference in the culture to qPCR ratio arising from uneven loading of culture and qPCR targets to recreational sites.

In Figure 8, paired *Enterococcus* densities for two New Jersey beaches, in Monmouth County (red circles) and Ocean County (blue circles), are illustrated. A dashed green regression line shows the best fit of a linear model to the raw data (not log-transformed) while a dashed red regression line shows a linear regression fit to the log-transformed data. The black 45° line is a hypothetical line showing perfect correspondence between *Enterococcus* densities when measured via the two methods. The paired data for these two beaches are similar to those reported for other beaches—high variability and uncertainty in the qPCR counts at low density, with a correlation approaching a 1:1 ratio at higher indicator density. In this example, it is not obvious whether the data from the two beaches are sufficiently similar to be pooled and used to

develop a single regression model for linking the indicator/method combinations. It is also not clear which water quality model is the best for fitting the data, though both have the potential for application. These two observations (i.e., uncertainty regarding data pooling and uncertainty regarding model form) indicate that statistical tests should be used to evaluate the similarity between datasets prior to their inclusion in the model, and that alternative functional forms should be evaluated for the model relating *Enterococcus* density as measured by culture (MF) methods to *Enterococcus* density as measured by qPCR.

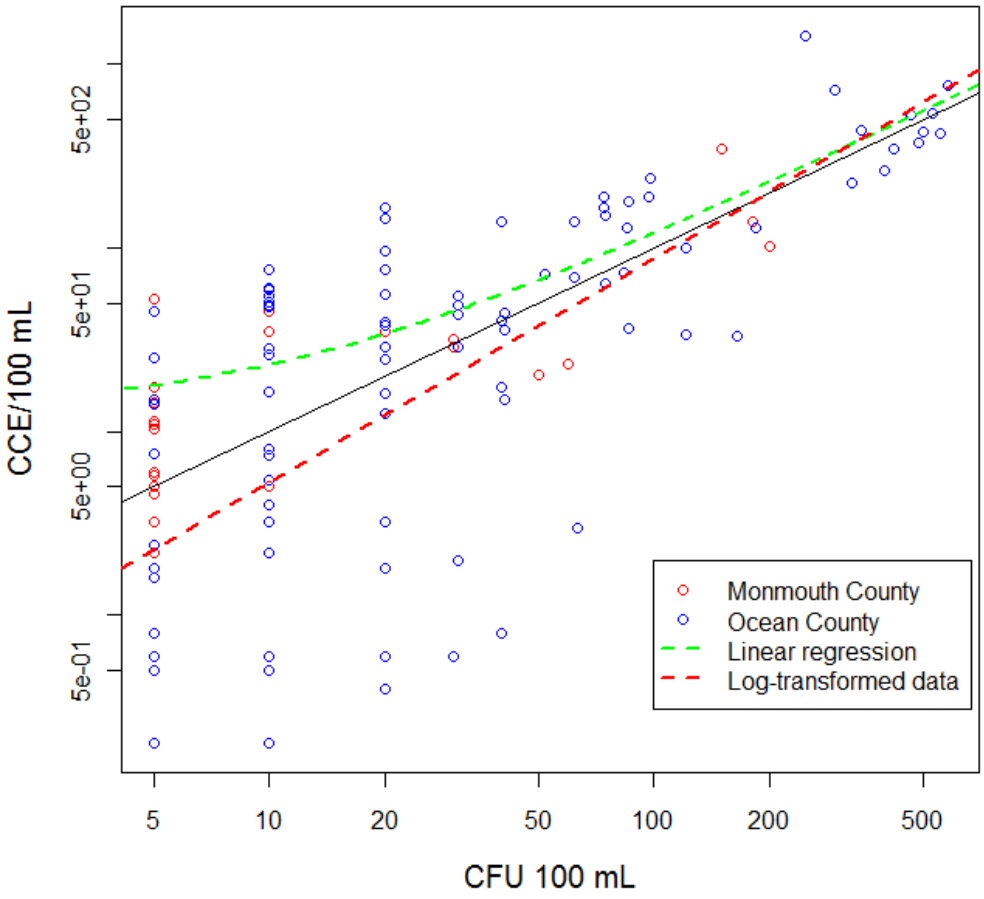

Figure 8. Typical plot of paired *Enterococcus* densities as measured by qPCR and culture-based methods

ANCOVA, a statistical method for assessing differences in model slopes and intercepts for data corresponding to different factors (e.g., different beaches or different times of day), can be used to assess whether trends observed at different locations or times are similar, and to help ensure that only related data are used in model development. In the event that paired water quality data are determined to be similar, the data may be pooled and model parameters may be estimated using the pooled dataset. If models using datasets that are determined to be dissimilar, it is assumed that there are differences in the systems that generated the data and those differences are not accounted for in the model relating the two datasets. The use of ANCOVA for developing datasets is illustrated with the following example.

As part of EPA's NEEAR studies conducted in the Great Lakes, *Enterococcus* densities as measured by MF and by qPCR were collected (Wade et al. 2006, 2008). Samples were collected three times per day (8:00 AM, 11:00 AM, and 3:00 PM), at multiple depths (waist, shin, and knee), and along three transects at each of four beaches studied. A model relating the paired *Enterococcus* densities as measured by both qPCR and MF can be developed from all the sites and times, or from data that are selected because they appear to have similar underlying trends. All the paired water quality data collected are shown in Figure 9, along with a 45° line showing perfect agreement. Although one may observe a general trend at high indicator densities, no trend is apparent at densities below roughly 500 CFU/100 mL. Additionally, there is no obvious choice for the form (equation relating *Enterococcus* density as measured by qPCR and as measured by membrane filtration) of a model relating the data from the two analytical methods. Figure 10 shows paired water quality data from Huntington Beach only, one of the four Great Lakes beaches studied. Again, the FIB densities appear linearly related at high densities and scattered at low densities, though the scatter is considerably less than that observed when data for all the beaches are plotted together.

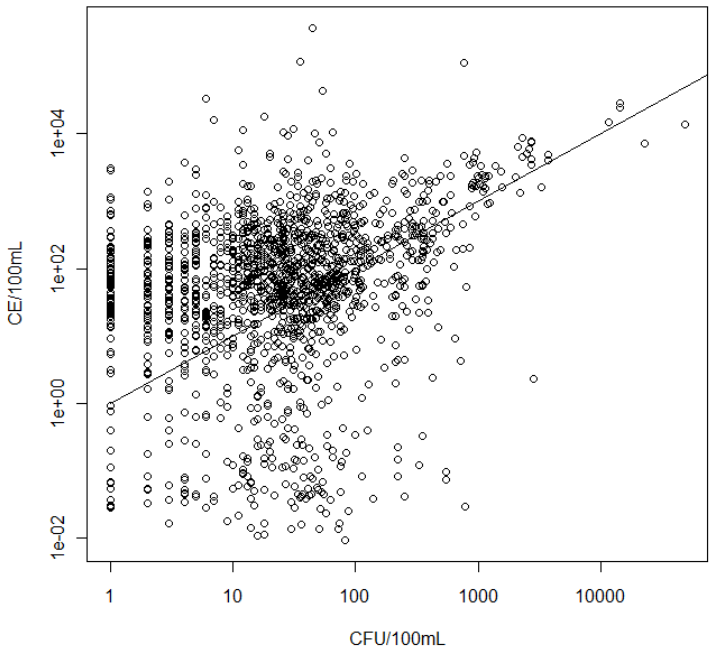

Figure 9. Paired qPCR and culture *Enterococcus* data, all NEEAR study Great Lakes beaches

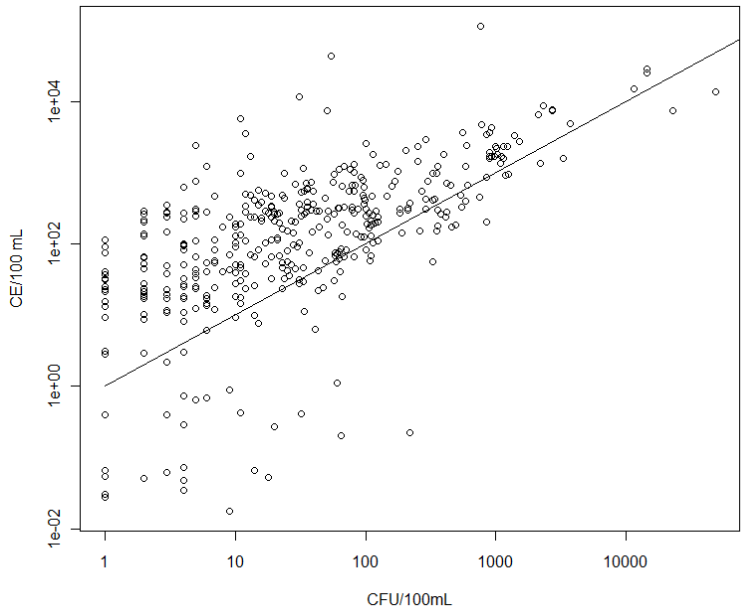

Figure 10. Paired culture and qPCR *Enterococcus* data, all samples from Huntington Beach

Figure 11 and Figure 12 show data for only 8:00 AM and 3:00 PM, respectively, along with 45 ° lines and regression fits developed via linear regression of the log-transformed densities. Note that the trend lines for the 8:00 AM and 3:00 PM samples are markedly different and that segregation of data by time of the day markedly reduces the scatter in the plots.

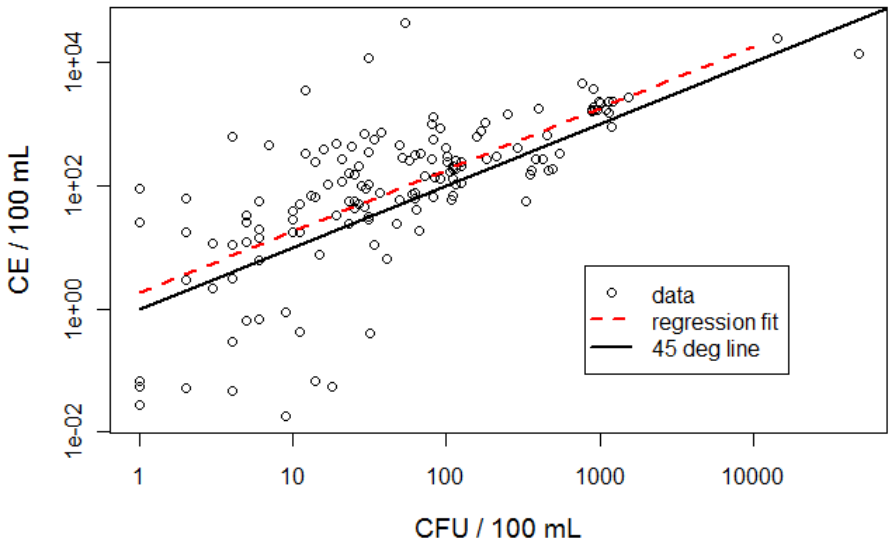

Figure 11. Paired qPCR and culture *Enterococcus* data, Huntington Beach: 8:00 AM samples only

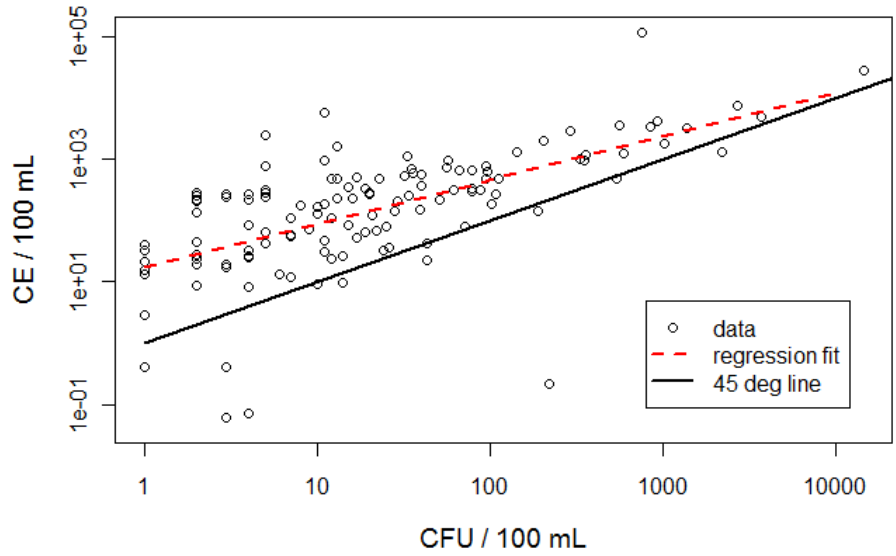

Figure 12. Paired qPCR and culture *Enterococcus* data, Huntington Beach: 3:00 PM samples only

ANCOVAs can be used to assess whether the differences in the regression lines observed for the 8:00 AM and 3:00 PM data are statistically significant. ANCOVAs were run to compare the regression lines for sets of data corresponding to each of the collection times at Huntington Beach. As shown in Table 10, the ANCOVA indicates that there are significant differences in the models for the 8:00 AM and 11:00 AM dataset, and for the 8:00 AM and 3:00 PM datasets, but not for the 11:00 AM and 3:00 PM datasets. This finding is not surprising given that the 8:00 AM and 3:00 PM samples do not correspond to the same model, and because solar radiation is known to reduce culture counts for afternoon samples while qPCR counts remain relatively constant throughout the day.

Table 10. Results of ANCOVAs for sets of data for Huntington Beach with different collection times

Collection times	P-value	
	Intercepts	Slopes
8:00 AM and 11:00 AM	<0.0001	0.002
8:00 AM and 3:00 PM	<0.0001	0.01
11:00 AM and 3:00 PM	0.52	0.60

Additional ANCOVAs were performed to compare models for all of the NEEAR freshwater study datasets and to determine whether sample water depth, time of day, and beach location altered the model relating qPCR and culture-based datasets. Findings from the ANCOVAs are as follows:

- Samples corresponding to different sample collection depths can be pooled.

- Models for different times of day differ for two beaches (Huntington and West Beach) but do not differ for the other two beaches (Silver Beach and Washington Park Beach). Agreement between models for different times of day for Silver Beach and Washington Park Beach might be the result of extreme scatter in the data for those two beaches. Based on these findings and knowledge of the variation in culture and qPCR signals during the day, 8:00 AM *Enterococcus* measured by MF are fit by different models than samples from other times of day..
- For these data, models for different beaches are dissimilar and pooling of the data may be inconsistent with simple linear models.

3.2.3 Demonstration of the Water Quality Linkage Approach using Linear Regression Models

In this section, the Water Quality Link approach is demonstrated using NEEAR study Great Lakes *Enterococcus* density data (as measured by both MF and qPCR), the NEEAR study freshwater health effects curve (presented in equation form in Table 5), and linear models relating log-transformed data of *Enterococcus* density as measured by qPCR and MF. As noted previously, these linear models may not adequately characterize the correlation between qPCR and culture *Enterococcus* density for some datasets. Therefore, the illustration presented below may be used for developing an understanding of the Water Quality Linkage approach but not for developing specific equivalences between densities of *Enterococcus* as measured by MF and as measured by qPCR. The Water Quality Link demonstration entails

1. selection of an *Enterococcus* density based on the NEEAR study freshwater beaches health effects curve;

2. use of linear regression to develop several candidate models relating density of *Enterococcus* as measured by qPCR to *Enterococcus* density as measured by MF; and

3. use of the linear relationships developed in step 2 to establish candidate *Enterococcus* densities as measured by MF to the *Enterococcus* density selected in step 1.

After the Water Quality Linkage approach is demonstrated the results are critically evaluated and used to suggest alternative Water Quality Linkage approaches.

The NEEAR study health effects curve for Great Lakes beaches (see Table 5) indicates that an *Enterococcus* density of about 125 qPCR CE/100 mL relates to a risk of about 45 NGI illnesses in 1000 swimmers. The rate 45 NGI illnesses in 1000 is approximately equivalent to a risk of 8 HCGI illnesses in 1000 swimmers—the level at which current (1986) freshwater criteria are based (see Section 3.1.4). Thus for this demonstration, the qPCR level of interest is selected to be 125 qPCR CE. This selection does not imply that new or revised criteria will be based on the Great Lakes health effects curve or on a risk level of 45 NGI illnesses in 1000 swimmers. It is selected for demonstration purposes only and because it is based on risk in a range consistent with current (USEPA 1986) criteria.

The choice of a model relating the *Enterococcus* density data from qPCR to density data from MF has a profound influence on the outcome of a Water Quality Linkage analysis. This is illustrated using the NEEAR study data. When data from all the NEEAR study freshwater

beaches are used to develop a model relating *Enterococcus* qPCR and culture densities (the purple line in Figure 13), the regression line has a slope <1 and falls significantly below the 45° line, for the culture-based density above about 100 CFU/100 mL. The model mean value for CFU equivalent to CE for 45 illnesses/1000 swimmers is 1212 CFU/100 mL and the confidence interval for CFU equivalent to CE for 45 illnesses/1000 swimmers is <519 CFU/100 mL, 4628 CFU/100mL>. Thus, these *Enterococcus* densities, measured by MF, are clearly very high and inconsistent with the selected level of tolerable risk.

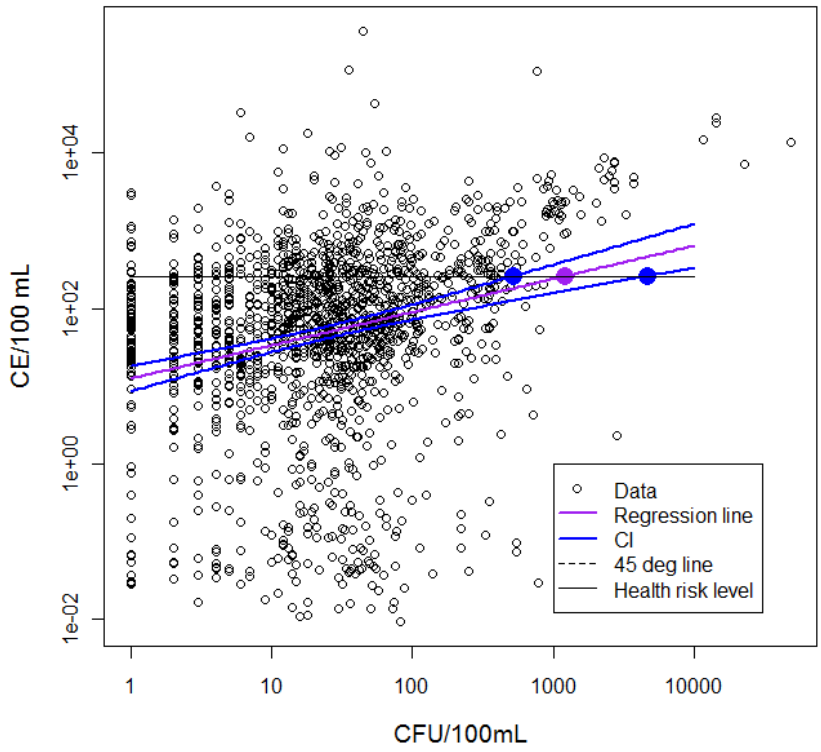

Figure 13. Linear model, data from all NEEAR freshwater beaches

When only Huntington Beach data are used, the results appear to be more reasonable. The model line (the purple line in Figure 14) still has a slope <1, which is inconsistent with expected trends at high indicator density, but the slope is steeper than that of the model for data from all beaches. The model mean value for CFU equivalent to CE for 45 illnesses/1000 swimmers is 10.8 CFU/100 mL, and the confidence interval around that value is <7.1 CFU/100mL, 15.4 CFU/100mL>. These equivalent values appear unrealistically low, though more reasonable than the equivalent values derived using the model for data from all beaches.

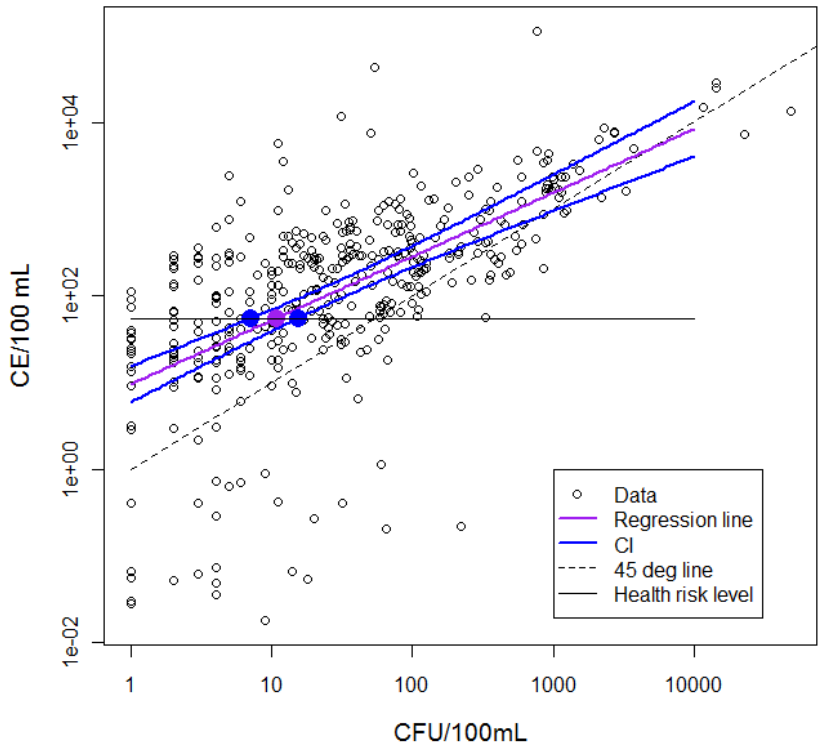

Figure 14. Linear model, data from Huntington Beach

The most reasonable equivalent culture densities result from a model developed using only 8:00 AM data from Huntington Beach. The model, equivalent value, and confidence interval are all shown in Figure 15. The model mean value for CFU equivalent to CE for 45 illnesses/1000 swimmers is 29.4 CFU/100 mL and the confidence interval around that value is <17.8 CFU/100mL, 45.9 CFU/100mL>.

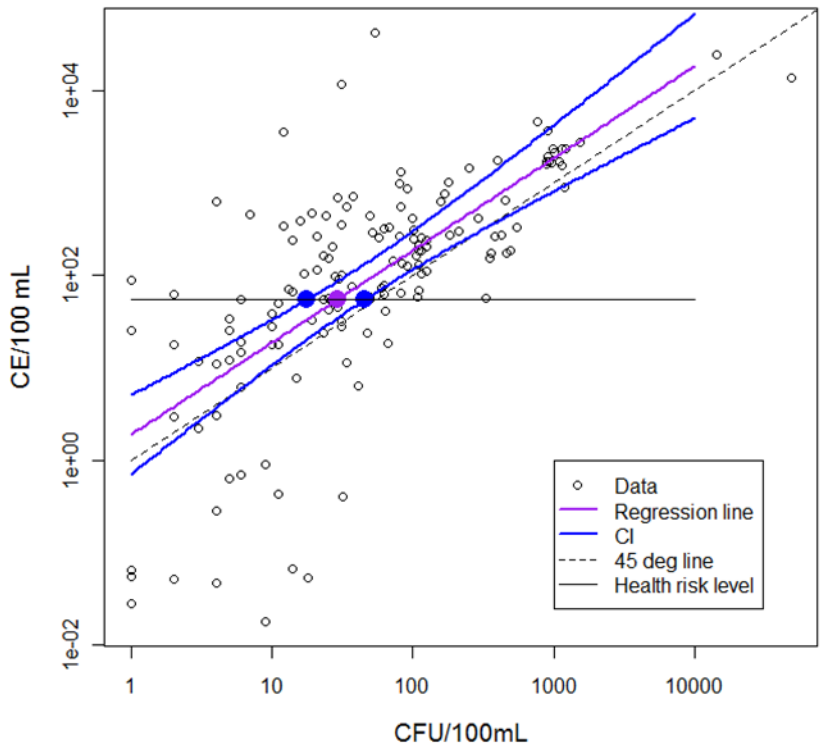

Figure 15. Linear model, 8:00 AM data from Huntington Beach

Based on the analysis of paired water quality NEEAR study *Enterococcus* density data measured by qPCR and MF and using simple linear model of log-transformed variables, several observations may be made, though the extent to which these may apply to datasets other than the ones studied is unknown. First, model selection (inclusive of data selection) profoundly impacts the Water Quality Linkage approach. For this dataset, when all data are pooled and a water quality model is generated, the resulting densities of *Enterococcus* as measured by MF that are equivalent to the density of *Enterococcus* as measured by qPCR appear very high in comparison to the current criteria. When only data from morning samples at a single beach are used to develop the linear model, the resulting equivalent *Enterococcus* densities for MF are closer to the current criteria. The non-robust performance of linear models of log-transformed indicator densities in this demonstration suggests exploration of alternative models that may be capable of modeling more complex phenomena and providing an adequate fit to a larger set of data.

3.2.4 Broken Stick Models of Log-Transformed Indicator Data

The relationship between paired *Enterococcus* densities as measured by MF and *Enterococcus* densities as measured by qPCR data might not be fit best by a simple linear regression model of log-transformed data. An alternative model evaluated in the development of the Water Quality Linkage approach is the broken stick regression model, also known as the segmented model (Draper and Smith 1998). In broken stick regression models, the data are divided in two parts, each of which is fit with a different linear model. The two linear models meet at a "kink" in the regression line. Broken stick regression models have four parameters (two slopes and the coordinates of the point where the segments of the broken stick meet) that may be determined via

optimization, whereas simple linear regression models have two parameters (slope and intercept). The broken stick model appears a good choice for analyzing paired densities of *Enterococcus* as measured by MF and *Enterococcus* as measured by qPCR because the data appear to behave differently at low and high indicator densities.

Broken stick regression models were developed for each of the four NEEAR Great Lakes study beaches separately. Two models were developed for each beach—one for all data for that beach pooled and one for only data from samples collected at 8:00 AM. The resulting curves are shown in Figure 16 to Figure 19.

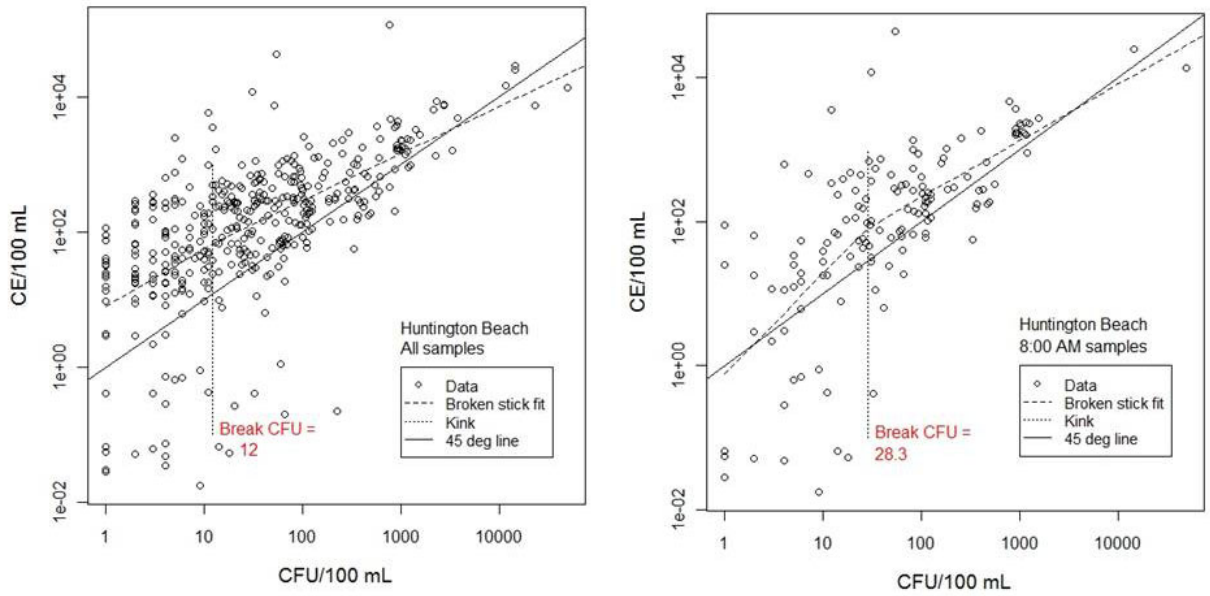

Figure 16. Huntington Beach broken stick model fits for all data and 8:00 AM data

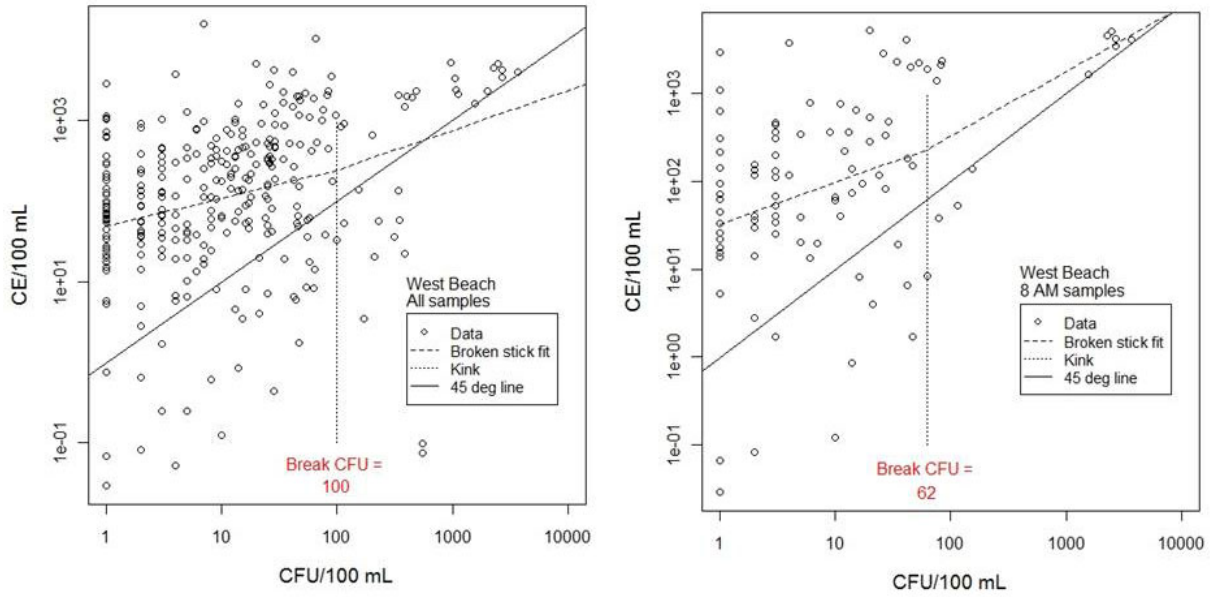

Figure 17. West Beach broken stick model fits for all data and 8:00 AM data

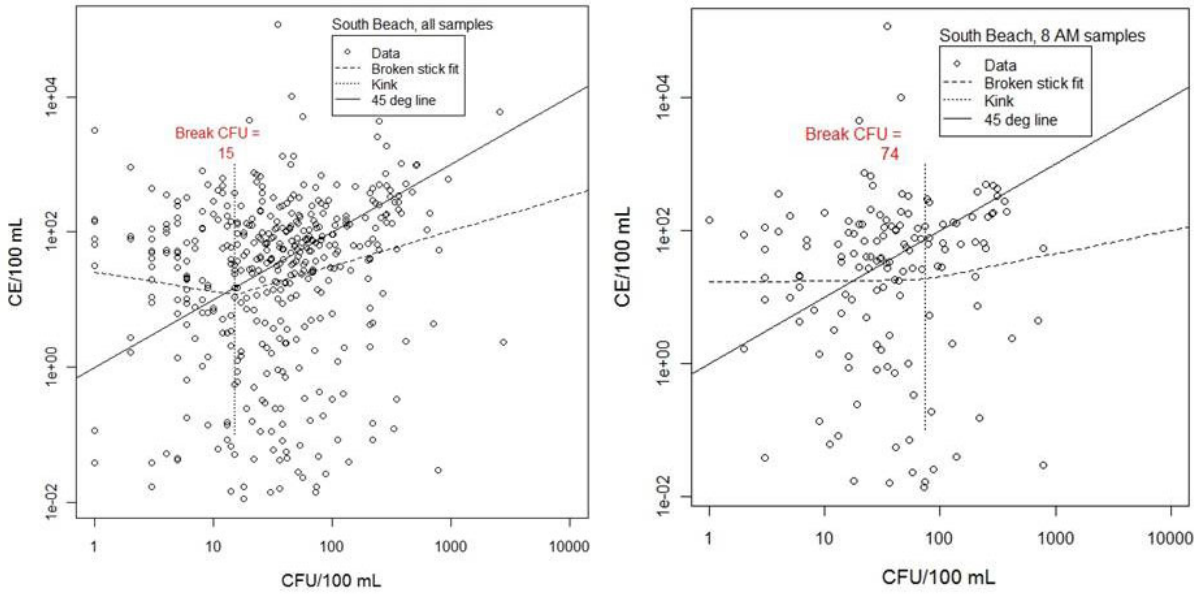

Figure 18. Silver Beach broken stick model fits for all data and 8:00 AM data

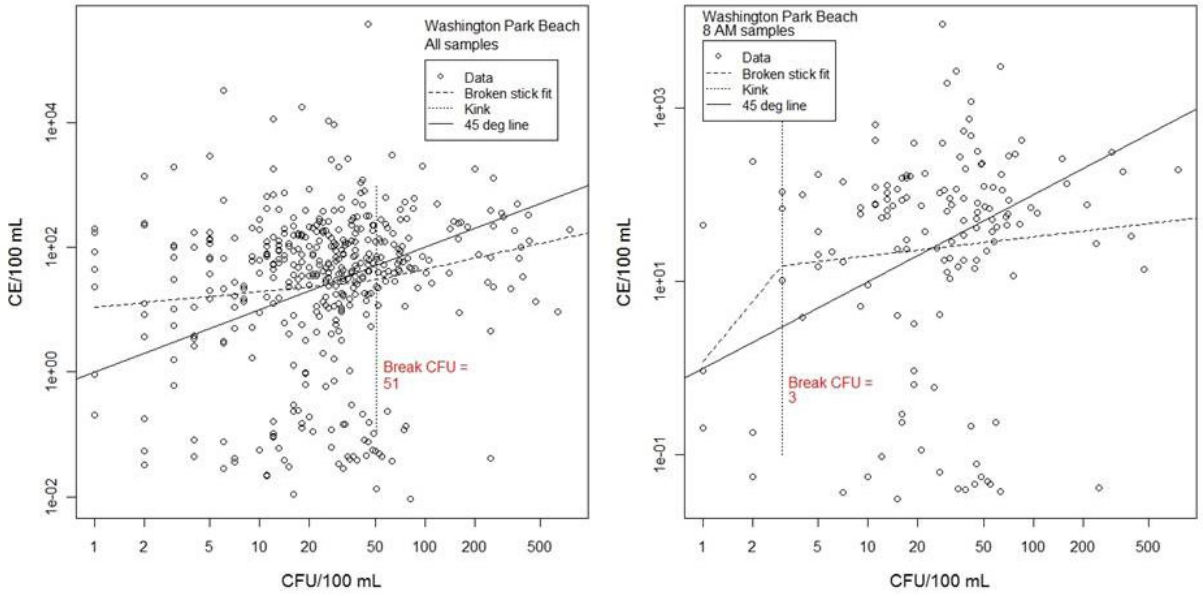

Figure 19. Washington Park Beach broken stick model fits for all data and 8:00 AM data

Broken stick regression models of Huntington and West Beach data both exhibit slopes closer to 1.0 in the high FIB density portion of the broken stick, compared to the low density portion. For the Washington Park and Silver Beach datasets, the presence of water quality data pairs with relatively high *Enterococcus* culture density and very low *Enterococcus* qPCR density appear to exert an undue influence on the slope portion of the broken stick in the high culture density data range. For the four beaches, the location of the "kink" in the stick varies within the range 12–100 CFU/100 mL for models of data including all sample times, and in the range 28–62 CFU/100 mL for the 8:00 AM samples of the Huntington and West Beach data.

Given the empirical and theoretical support for different correlation of qPCR and culture counts at high and low densities, the broken stick model offers advantages over models of simple linear regression of log-transformed densities, including

- better fits to data,
- reduction in heteroskedasticity of regression fits, and
- consistency with observed relationships between qPCR and culture densities.

Regression analysis of correlations between the qPCR and culture densities, as well as the literature, support different relationships between the two water quality measures at low and high indicator counts. At low indicator densities, the relationship of the culture indicator with the fecal pollution source becomes tenuous and the qPCR counts become highly uncertain. At higher densities, the abundance in targets for the two methods is correlated. These findings both support and suggest performing regression analysis only for those data pairs for which culture densities exceed a threshold value.

Three regression analyses were performed to evaluate whether using only data above such threshold values yielded robust models for estimating culture densities that are equivalent to qPCR densities. In all of the analyses, only FIB density data collected in the morning (8:00 AM

sample time) were included, due to potential discrepancies between enumeration methods at later times. In the first analysis, only points above the "kink" in the stick were retained for analysis for each of the four Great Lakes beaches at which EPA's NEEAR epidemiology studies were conducted. The data retained for each beach were determined by the "kink" location for each of the four beaches. In the second and third analyses, a threshold common to all beaches was chosen as the basis for retaining data. Two thresholds were explored—80 CFU/100mL and 104 CFU/100 mL. The value 104 CFU/100 mL was selected based on the current single sample maximum for designated beaches; 80 CFU/100 mL was selected because it lies between the current criteria and the geometric mean on which it is based (35 CFU/100 mL).

Plots showing the resulting three regression models and the culture equivalent to a qPCR density of 126 CE/100 mL are shown in Figure 20 (data above the kink in the stick), Figure 21 (data above 80 CFU/100 mL), and Figure 22 (data above 104 CFU/100 mL). The qPCR density chosen for this illustration (126 CE/100 mL) is a plausible value for qPCR-based criteria (Table 9). In all cases, the new regression models and qPCR value of interest result in equivalent culture criteria exceeding the current single sample maximum value for infrequently-used fresh waters of 151 CFU enterococci / 100 mL. Regression model lines and culture equivalents to a potential qPCR criterion of 126 CE/100 mL are provided in Table 11.

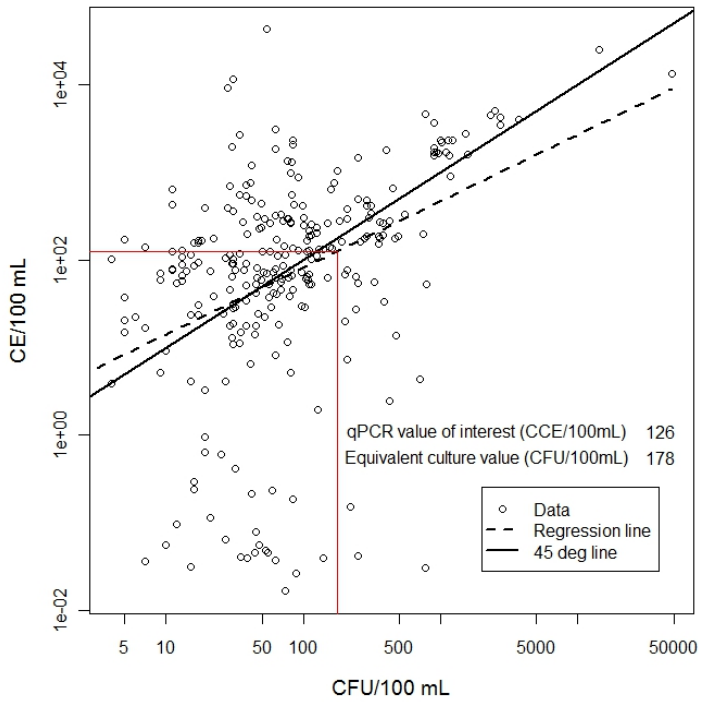

Figure 20. Model resulting from retention of all data above the "kink" in the broken stick

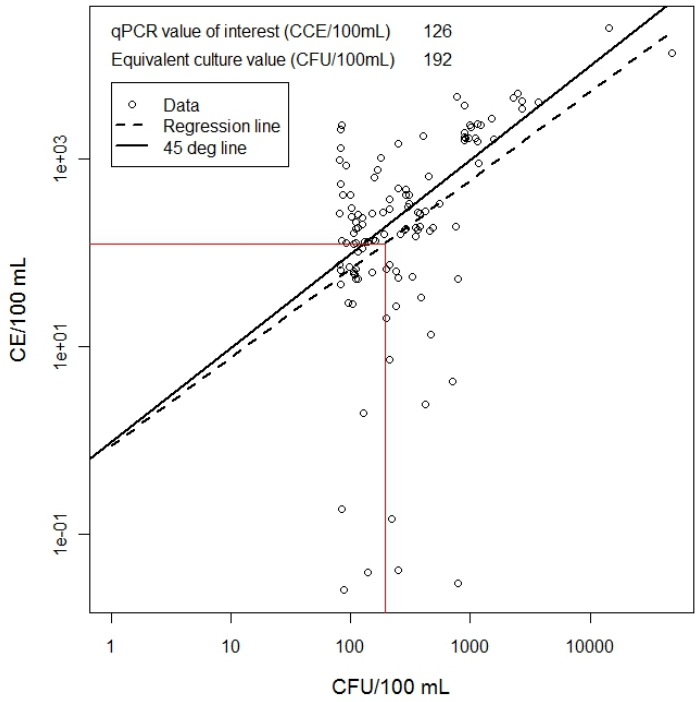

Figure 21. Model resulting from retention of all data above 80 CFU/100 mL

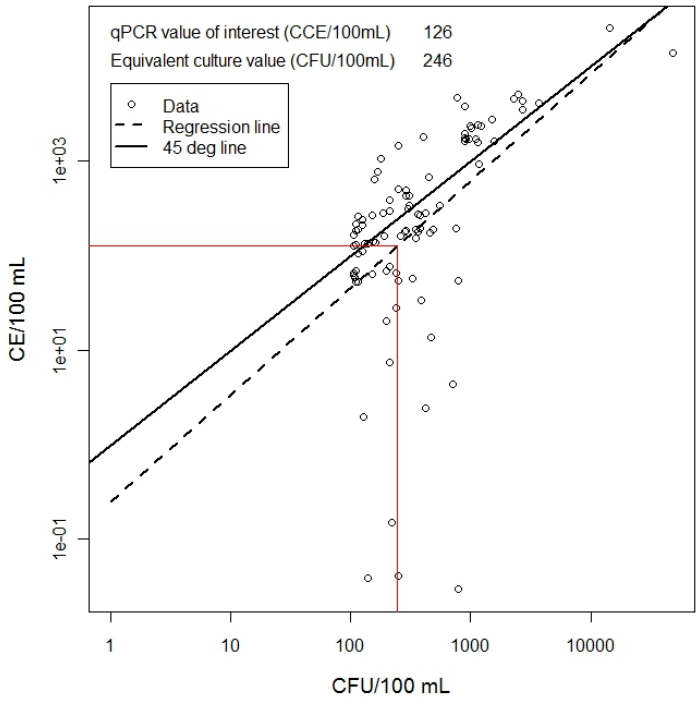

Figure 22. Model resulting from retention of all data above 104 CFU/100 mL

Table 11. Summary of models resulting from retention of data pairs with culture counts above three thresholds

Treatment	Regression line	Equivalent culture density of interest (CFU/100 mL)
Retain data for which CFU count is above the break in the stick	$\log_{10}(C_{qPCR}) = 0.393 + 0.759 \log_{10}(C_{CFU})$	178
Retain data for which CFU count is above 80 CFU/100 mL	$\log_{10}(C_{qPCR}) = -0.056 + 0.944 \log_{10}(C_{CFU})$	192
Retain data for which CFU count is above 126 CFU/100 mL	$\log_{10}(C_{qPCR}) = -1.289 + 1.358 \log_{10}(C_{CFU})$	246

Several observations can be made regarding development of a regression model based on data only above thresholds. First, the model results appear quite sensitive to the selection of the break point above which data are retained. In this regard, there does not appear to be a widely accepted methodology for selecting the appropriate break point. Both Lavender and Kinzelman (2009) and Byappanahalli et al. (2010) determined that the qPCR-culture relationship varies with environmental conditions which, in turn, vary differently at different sites. Second, in all cases, data pairs with much lower qPCR densities than culture densities appear to have a pronounced effect on the regression models. At present, we have no explanation for the occurrence of those data pairs. Because the pairs occur above the cutoff value for culture densities, it is unlikely that sampling error (i.e., small qPCR sample size in terms of original sample volume) is the cause. In summary, regression models based on the broken stick or threshold approach were dependent on a cut-off value, the choice of which could be viewed as arbitrary. Further, culture densities equivalent to potential qPCR densities were higher than the current single sample maximum for infrequently-used freshwater beaches when the Water Quality Link approach is applied to the NEEAR study freshwater indicator dataset, using a broken stick regression model to relate densities of *Enterococcus* as measured by MF and qPCR.

4 Discussion

Two approaches for linking indicator-method combinations to a common risk level were demonstrated—the Risk Link approach and the Water Quality Link approach. Using currently available data and techniques, the exploration of these methods meets both the objective of developing quantifiable relationships between indicator-method combinations and provides practical experience in the complexities underlying the two approaches.

Applying the Risk Link approach can be relatively straightforward when multiple indicator health relationships data are available from a single epidemiology study, as demonstrated in the *Bacteroidales* (measured by qPCR) and *Enterococcus* (measured by qPCR) Risk Link comparison. However, when directly comparable data are not available, additional steps are added to the Risk Link approach. These steps may include, but are not limited to, the harmonization of GI definitions used in two epidemiology studies or the generality of health effects relations established for sites with specific sources of fecal pollution.

Regarding the Water Quality Link approach, generating a relationship between indicator densities, as measured by qPCR and culturable methods, is more complex than simple linear or broken stick regression models. In our demonstration, the relationship, between paired water quality data corresponding to different indicator-method combinations, appears to vary from beach to beach. The Water Quality Link may be a useful tool for the development of site-specific standards by states; but as formulated in this report, is not necessarily useful for the development of National criteria.

It is possible that additional analyses or datasets might allow for alternative quantifiable relationships to be explored using both Link approaches. For example, if statistical techniques are developed that allow for the direct comparison of results between RCT and PC epidemiology study designs, additional health effects relationships, such as those from the Epibathe studies, could be evaluated.

5 References

Abdelzaher, A., M. Wright, C. Ortega, H. Solo-Gabriele, G. Miller, S. Elmir, X. Newman, P. Shih, T.D. Bonilla, C.J. Palmer, T. Scott, J. Lukasik, V.J. Harwood, S. McQuaig, C. Sinigalliano, M. Gidley, L. Plano, X. Zhu, J.D. Wang, and L. Fleming. 2010. Presence of pathogens and indicator microbes at a non-point source subtropical recreational marine beach. Applied and Environmental Microbiology 76(3): 724-732.

Agudel, R.M., Codony, F., Adrado, B., Fittipaldi M., Peñuela, G., and J. Morató. 2010 Monitoring bacterial faecal contamination in waters using multiplex real-time PCR assay for *Bacteroides* spp. and faecal enterococci. Water SA, 36(1): 127-132.

Ahmed, W., S. Sawant, F. Huygens, A. Goonetilleke, and T. Gardner. 2009. Prevalence and occurrence of zoonotic bacterial pathogens in surface waters determined by quantitative PCR. Water Research 43: 4918-4928.

Bae, S., and S. Wuertz 2009. Discrimination of viable and dead fecal bacteroidales bacteria by quantitative PCR with propidium monoazide. Applied and Environmental Microbiology 75(9): 2940-2944.

Bolster, C.H., J.M. Bromley, and S.H. Jones. 2005. Recovery of chlorine-exposed *Escherichia coli* in estuarine microcosms. Environmental Science and Technology 39(9):3083-3089

Bower, P.A., C.O. Scopel, E.T. Jensen, M.M. Depas, and S.L. McLellan. 2005. Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. Applied and Environmental Microbiology 71(12): 8305-8313.

Bremer, H. and Dennis, P. P. 1996. Modulation of chemical composition and other parameters of the cell growth rate. In *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, 2nd ed., pp. 1553–1568. Edited by F. C. Neidhardt and others. Washington DC: American Society for Microbiology.

Byappanahalli, M.N., R.L. Whitman, D.A. Shively, and M.B. Nevers. 2010. Linking non-culturable (qPCR) and culturable enterococci densities with hydrometeorological conditions. Science of the Total Environment 408: 3096-3101.

Calderon, R.L., E.W. Mood, and A.P. Dufour . 1991. Health effects of swimmers and nonpoint sources of contaminated water. International of Envionmental Health Research 1: 21-31.

Colford, J.M.J., T.J. Wade, K.C. Schiff, C.C. Wright, J.F. Griffith, S.K. Sandhu, S. Burns, M. Sobsey, G. Lovelace, and S.B. Weisberg. 2007. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiology 18(1): 27-35.

Converse, R., A. Blackwood, M. Kirs, J. Griffith, and R. Noble. 2009. Rapid QPCR-based assay for fecal *Bacteroides* spp. as a tool for assessing fecal contamination in recreational waters. Water Research 43(19): 4828-4837.

Draper, N.R., and Smith, H. 1998. *Applied Regression Analysis,* 3rd ed. John Wiley and Sons, NY.

Duprey, E., M.P. Caprais, A. Derrien, and P. Fach, 1997. *Salmonella* DNA persistence in natural seawaters using PCR analysis. Journal of Applied Microbiology 82: 507-510.

European Commission (Epibathe consortium). 2009a. Accessible Report to Commission Staff Responsible for Reporting to the Parliament and to the Council Under Article 14 of the Bathing Water Directive of 15th February 2006. Assessment of Human Health Effects

Caused by Bathing Waters (EpiBathe project). European Union Sixth Framework Programme, Scientific Support to Policies, FP6 Project Reference 022618.

European Commission (Epibathe consortium). 2009b. Appendix 2 Section 1. Risk of Gastroenteritis from Recreational Bathing Water: Analysis of Epibathe data from Hungary and Spain.

Elmir S.M., T. Shibata, H.M. Solo Gabriele, C.D. Sinigalliano, M.L. Gidley, G. Miller, L.R.W. Plano, J. Kish, K. Withum, and L.E. Fleming. 2009. Quantitative evaluation of enterococci and *Bacteroidales* released by adults and toddlers in marine water. Water Research 43(18): 4610-4616.

Ferretti, J.A., Tran, H.V., Loftin, V.K., Cosgrove, B., and Protonentis, J. 2008. Comparison of *Enterococcus* Measurements in Marine Beach and Bay Samples by Quantitative (Real-Time) Polymerase Chain Reaction, Membrane Filtration, and Enterolert. Edison, NJ: US EPA Region 2.

Fleisher, J.M., D. Kay, R.L. Salmon, F. Jones, M.D. Wyer, and A.F. Godfree. 1996. Marine waters contaminated with domestic sewage: nonenteric illnesses associated with bather exposure in the United Kingdom. American Journal of Public Health 86(9): 1228-1234.

Fleisher, J.M., L.E. Fleming, H.M. Solo-Gabriele, J.K. Kish, C.D. Sinigalliano, L. Plano, S.M. Elmir, J.D. Wang, K. Withum, T. Shibata, M.L. Gidley, A. Abdelzaher, G. He, C. Ortega, X. Zhu, M. Wright, J. Hollenbeck, and L.C. Backer. 2010. The BEACHES Study: health effects and exposures from non-point source microbial contaminants in subtropical recreational marine waters. International Journal of Epidemiology 39(5): 1291-1298.

Griffith, J., K.C. Schiff, G.S. Lyon, and J.A. Fuhrman. 2010. Microbiological water quality at non-human influenced reference beaches in southern California during wet weather. Marine Pollution Bulletin 60(4): 500-508.

Haile, R.W., J.S. Witte, M. Gold, R. Cressey, C. McGee, R.C. Millikan, A. Glasser, N. Harawa, C. Ervin, P. Harmon, J. Harper, J. Dermand, J. Alamillo, K. Barrett, M. Nides, and G. Wang. 1999. The health effects of swimming in ocean water contaminated by storm drain runoff. Epidemiology 10(4): 355-363.

Haugland, R.A., S.C. Siefring, L.J. Wymer, K.P. Brenner, and A.P. Dufour. 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.

He, J.-W., and S. Jiang. 2005. Quantification of enterococci and human adenoviruses in environmental samples by real-time PCR. Applied and Environmental Microbiology 71(5): 2250-2255.

Khan, I.U.H., V. Gannon, R. Kent, W. Koning, D.R. Lapen, J. Miller, N. Neumann, R. Phillips, W. Robertson, E. Topp, E. van Bochove, and T.A. Edge. 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.

Kinzelman, J.L., R.N. Bushon, R.T. Noble, and S.E. Dorevitch. 2010. WERF - PHASE I Progress Report. Comparative Evaluation of Molecular and Culture Methods for Fecal Indicator Bacteria for use in Inland Recreational Waters. In: WERF.

Lavender, J.S., and J.L. Kinzelman. 2009. A cross comparison of QPCR to agar-based or defined substrate methods for determination of *Escherichia coli* and enterococci in municipal water quality monitoring programs. Water Research 43: 4967-4979.

Lleo, M.M., B. Bonato1, M.C. Tafi1, C. Signoretto1, C. Pruzzo2 and P. Canepari. 2005. Molecular vs culture methods for the detection of bacterial faecal indicators in groundwater for human use Letters in Applied Microbiology 40:289–294

Marion, J.W., J. Lee, S. Lemeshow, and T.J. Buckley. 2010. Association of illness and recreational water exposure during advisory and non-advisory conditions at an inland U.S. beach. Water Research 44(16): 4796-4804.

Morrison, C.R., D.S. Bachoon, and K.W. Gates 2008. Quantification of enterococci and bifidobacteria in Georgia estuaries using conventional and molecular methods. Water Research 42(14): 4001-4009.

Nocker, A., and A.K. Camper. 2006. Selective removal of DNA from dead cells of mixed bacterial communities by use of ethidium monoazide. Applied and Environmental Microbiology 72(3): 1997-2004.

Nocker, A., C.Y. Cheung, and A.K. Camper. 2006. Comparison of propidium monoazide with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal of DNA from dead cells. Journal of Microbiological Methods 67: 310-320.

NRC (National Research Council). 2004. *Indicators for Waterborne Pathogens*. Washington DC: National Academies Press.

Ortega, C., M. Helena, S. Gabriele, A. Abdelzaher, M. Wright, Y. Deng, and L.M. Stark 2009. Correlations between microbial indicators, pathogens, and environmental factors in a subtropical estuary. Marine Pollution Bulletin 58(9): 1374-1381.

Petterson, S.R., R.S. Signor, and N.J. Ashbolt. 2007. Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentrations for QMRA. Journal of Water and Health 5(Supplement 1): 51-65.

Prüss, A. 1998. Review of epidemiological studies on health effects from exposure to recreational water. International Journal of Epidemiology 27(1): 1-9.

Rothrock, M. J. J., K. L. Cook and K. H. Bolster 2009. Comparative quantification of *Campylobacter jejuni* from environmental samples using traditional and molecular techniques. Canadian Journal of Microbiology 55(6):633-641.

Siefring, S.C., M. Varma, E. Atikovic, L.J. Wymer, and R.A. Haugland. 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. Journal of Water and Health 6(2): 225-237.

Shanks, O., C.A. Kelty, M. Sivaganesan, M. Varma, and R. Haugland. 2009. Quantitative PCR for genetic markers of human fecal pollution. Applied and Environmental Microbiology 75(17): 5507-5513.

Sinigalliano, C.D., J.M. Fleisher, M.L. Gidley, H.M. Solo-Gabriele, T. Shibata, L. Plano, S.M. Elmir, D. Wanless, J. Bartkowiak, R. Boiteau, K. Withum, A.M. Abdelzaher, G. He, C. Ortega, X. Zhu, M.E. Wright, J. Kish, J. Hollenbeck, T. Scott, L.C. Backer, and L.E. Fleming. 2010. Traditional and molecular analyses for fecal indicator bacteria in nonpoint source subtropical recreational marine waters. Water Research 44: 3763-3772.

Sinton, L.W., C.H. Hall, P.A. Lynch, and R.J. Davies-Colley. 2002. Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Applied and Environmental Microbiology 68(3): 1122-1131.

USEPA (U.S. Environmental Protection Agency). 1983. *Health Effects Criteria for Marine Recreational Waters*. EPA-600/1-80-031. Research Triangle Park, NC: USEPA.

USEPA (U.S. Environmental Protection Agency). 1984. *Health Effects Criteria for Fresh Recreational Water*. EPA-600/1-84-004. Research Triangle Park, NC: USEPA.

USEPA (U.S. Environmental Protection Agency). 1986. *Ambient water quality for bacteria - 1986*. EPA440/5-84-002.

USEPA (U.S. Environmental Protection Agency). 2000. Improved Enumeration Methods for the Recreational Water Quality Indicators: Enterococci and *Escherichia coli*.

USEPA (U.S. Environmental Protection Agency). 2007. *Critical Path Science Plan for the Development of New or Revised Recreation Water Criteria*. Washington DC: Office of Water and Office of Research and Development, 823-R-08-002

Varma, M., R. Field, M. Stinson, B. Rukovets, L. Wymer, and R. Haughland. 2009. Quantitative real-time PCR analysis of total and propidium monozide-resistant fecal indicator bacteria in wastewater. Water Research (early publication online ahead of print).

Wade, T.J., N. Pai, J.N. Eisenberg, and J.M. Colford. 2003. Do U.S. Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. Environmental Health Perspectives 111(8): 1102-1109.

Wade, T.J., R.L. Calderon, E. Sams, M. Beach, K.P. Brenner, A.H. Williams, and A.P. Dufour. 2006. Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. Environmental Health Perspectives 114(1): 24-28.

Wade, T.J., R.L. Calderon, K.P. Brenner, E. Sams, M. Beach, R. Haugland, L. Wymer, and A.P. Dufour. 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., E. Sams, K.P Brenner, R. Haugland, E. Chern, M. Beach, L. Wymer, C.C. Rankin, D. Love, Q. Li, R. Noble, and A.P. Dufour. 2010. Rapidly measured indicators or recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. Environmental Health 9: 66. http://www.ehjournal.net/contents/9/1/66

Walters, S.P., K.M. Yamahara, and A.B. Boehm. 2009. Persistence of nucleic acid markers of health-relevant organisms in sewater microcosms. Water Research (early publication online ahead of print).

WERF (Water Environment Research Federation). 2009. *Report on the Experts Scientific Workshop on Critical Research and Science Needs for the Development of Recreational Water Quality Criteria for Inland Waters.* Dallas/Ft. Worth, Texas, February 18-20, 2009. PATH4W0. Alexandria, VA.

Whitman, R.L., Z. Ge, M.B. Nevers, A.B. Boehm, E.C. Chern, R.A. Haugland, A.M. Lukasik, M. Molina, K. Przybyla-Kelly, D.A. Shively, E.M. White, R.G. Zepp, and M.N. Byappanahalli. 2010. Relationship and variation of qPCR and culturable enterococci estimates in ambient surface waters are predictable. Environmental Science and Technology 44: 5049-5054.

WHO (World Health Organization). 2003. *Guidelines for Safe Recreational Water Environments. Volume 1 Coastal and Fresh Waters*. World Health Organization. Geneva, Switzerland.

Wiedenmann, A., P. Kruger, K. Dietz, J. Lopez-Pila, R. Szewzyk, and K. Botzenhart. 2006. A randomized controlled trial assessing infectious disease risks from bathing in fresh recreational waters in relation to the concentration of *Escherichia coli*, intestinal enterococci*, Clostridium perfringens*, and somatic coliphages. Environmental Health Perspectives 114:228-236.

Wymer, L., K.P. Brenner, J.W. Martinson, W.R. Stutts, S.A. Schaub, and A.P. Dufour 2005. *The EMPACT Beaches Project: Results from a Study on Microbiological Monitoring in Recreational Waters*. EPA 600/R-04/023 USEPA, Office of Research and Development, National Exposure Research Laboratory.

Yamahara, K.M., S.P. Walters, and A.B. Boehm. 2009. Growth of enterococci in unaltered, unseeded beach sands subjected to tidal wetting. Applied and Environmental Microbiology 75: 1517-1524.

Zmirou, D., L. Pena, M. Ledrans, and A. Letertre. 2003. Risks associated with the microbiological quality of bodies of fresh and marine water used for recreational purposes: summary estimates based on published epidemiological studies. Archives of Environmental Health 58(11): 703-771.