Journal of Shellfish Research, Vol. 28, No. 2, 263–271, 2009.

EVALUATION OF FECAL POLLUTION INDICATORS IN AN OYSTER QUALITY ASSURANCE PROGRAM: APPLICATION OF EPIDEMIOLOGICAL METHODS

DAMIAN M. OGBURN[1,*] AND IAN WHITE[2]

[1]*URS Australia Pty Ltd, Level 3, 20 Terrace Rd, East Perth, Western Australia, 6004;* [2]*Fenner School for Environment and Society, Australian National University, Canberra, ACT 0200, Australia*

ABSTRACT Shellfish Quality Assurance programs rely on the use of bacteriological fecal pollution indicators as a routine monitoring tool for risk assessment and management in shellfish growing areas. Internationally, shellfish programs vary in detail but are typically based on either the United States model, as in Australia, which relies on the enumeration of indicator bacteria in water as a measure of fecal pollution, or the European Union model, which assesses the exposure of production sites to fecal pollution by determining indicator bacteria present in shellfish flesh. The equivalency of these approaches for delivering the same level of public health protection is not immediately apparent Both indicator tests are used in combination in a very conservative approach adopted in the Australian state of New South Wales (NSW) by the Shellfish Program for the oyster (predominately Sydney rock oyster) fishery, the country's oldest and the state's most productive fishery. This study reports on the analysis of a large data set of these two indicators taken at multiple sites within three NSW estuaries between 2000 and 2005. We report on performance measures used routinely in epidemiological studies to compare the two tests. These measures demonstrate poor agreement between the two tests. When the harvest area is deemed closed by the water test the odds of a negative test based on shellfish flesh is approximately one in four. Conversely, the odds of a positive test based on shellfish flesh occurring during extended periods of dry weather when the water test is clear exceed one in five and are seemingly random in occurrence. We are able to demonstrate a highly significant statistical relationship between the water test and environmental covariates of rainfall, salinity, and water temperature. We use k-fold cross validation methods to develop predictive models using these environmental covariates that can account for over 90% of the variation in water test readings. In contrast we are unable to demonstrate a valid predictive model for the oyster flesh test using these covariate. We discuss the results of these analyses and suggest that there are significant issues with the efficacy of the shellfish flesh test and, in particular, the unexplained systematic error that occurs at a high rate in the current program. The dual use of these tests in harvest management results in routine uncertainty that must create significant unpredictability and costs in trade. In turn, this uncertainty and the costs of testing result in levels of commercial risk that could prove unsustainable. The study highlights the need for the performance of tests to be rigorously appraised in shellfish quality assurance programs. In so doing it may be possible to maintain public health standards while minimizing unnecessary disruptions and costs in the trade of fresh oysters.

KEY WORDS: oyster, shellfish quality assurance, fecal indicators, epidemiology, harvest

INTRODUCTION

The oyster industry in the eastern Australian state of New South Wales (NSW) is the major state fishery with over 3,000 hectares of leases in 38 estuaries along the entire coast (see Fig. 1). Oyster leases for farming predominantly the Sydney rock oyster, *Saccostrea glomerata* (Gould, 1850, were first granted in NSW in 1884 relatively close to the capital, Sydney, and since then suitable lease areas have been identified in other estuaries, mostly through industry trial and error. The majority of current lease areas are intertidal and have been in place for at least several decades (Ogburn et al. 2007).

Food-borne diseases associated with contaminated shellfish have, until relatively recently, only been considered as bacterial pathogens (Jackson & Ogburn 1999). Historically, bacterial diseases such as cholera and typhoid fever, associated with human fecal contamination of the environment were the chief concern. Shellfish also provide a potential vehicle of transmission for a variety of other infectious or otherwise injurious agents. These include natural constituents of the estuarine environment such as potentially toxic algal species (e.g., Hallegraeff 2002), and introduced agents such as human enteric viruses (Dorairaj & Miller 2001).

In the majority of cases, viruses found in aquatic environments that are infectious to humans are derived solely from human fecal waste. Whereas fecal waste from other animals does contain viruses, these are not generally pathogenic to humans (Dorairaj & Miller 2001). Human gastrointestinal (enteric) viruses are released with excreta, and exist in the environment as nonreplicating particles called virions. Waterborne viruses can only replicate by infecting a host cell with their genes, causing the host to generate more viruses (Dorairaj & Miller 2001). In recent years, international epidemiological evidence suggests that human enteric viruses, principally Noroviruses and hepatitis A virus, are now the most common aetiological agents transmitted by fresh bivalve shellfish (Lees 2000).

Internationally and in NSW, shellfish programs rely on fecal coliform bacteria in water and *Escherichia coli* in oyster meat as indicators of the quality of the growing waters and the end product, respectively. As a very conservative precaution, both indicator tests are currently used in the NSW Shellfish Program. The inadequacy of these fecal indicators to predict the presence of infectious viruses in marine waters is well documented (Henshilwood et al. 1998, Guyader et al. 1998, Griffin et al. 2001, Wetz et al. 2004). Development of molecular techniques for direct detection of enteric viruses in marine monitoring programs is continuing (Green & Lewis 1999, Guyader et al. 2000, Noble & Fuhrman 2001, Jiang et al. 2001). However, one problem with these molecular techniques is that a positive result may indicate the presence of nonviral particles rather than an actual health risk. There is currently no internationally accepted

*Corresponding author. E-mail: Damian_Ogburn@URSCorp.com

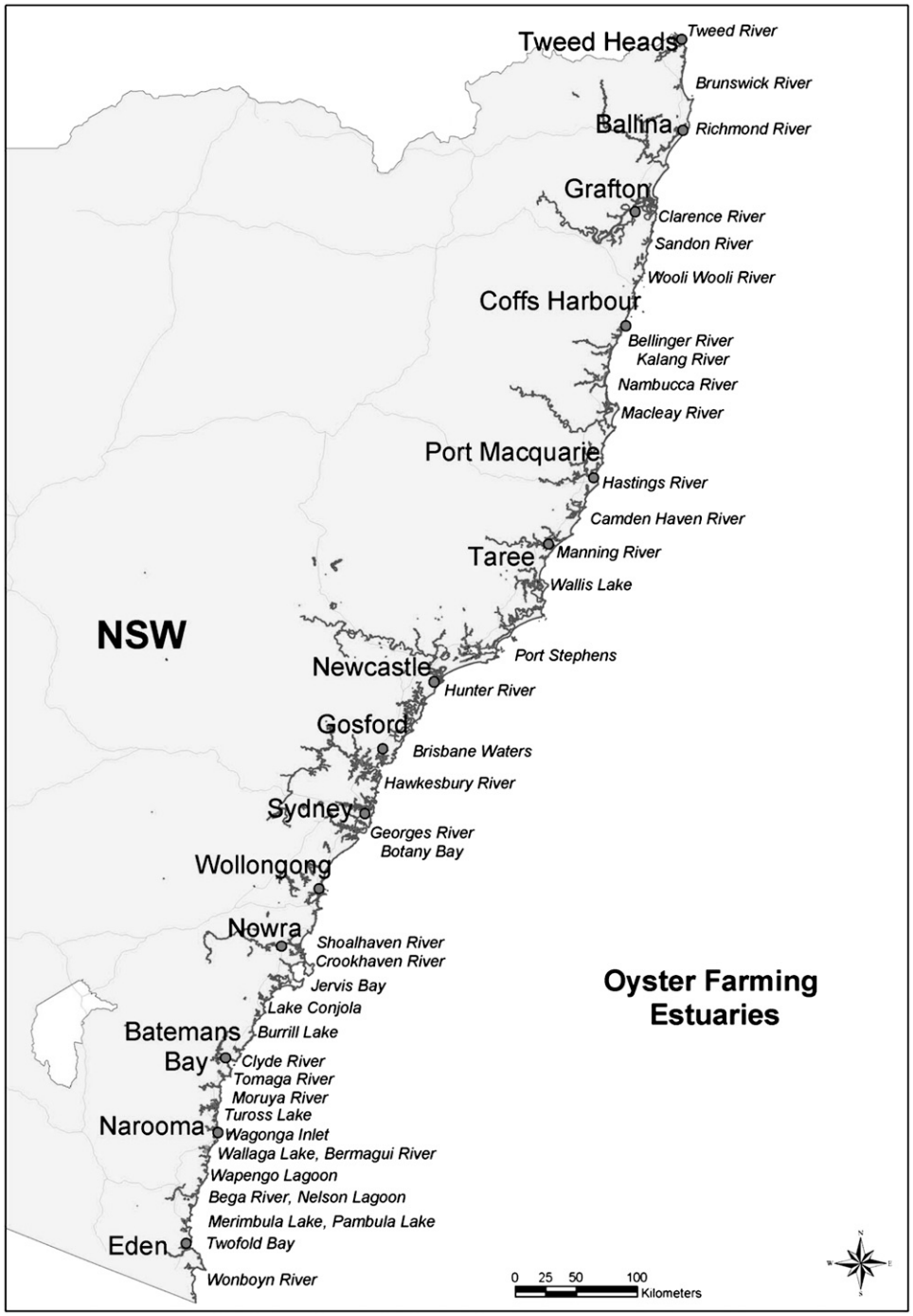

Figure 1. The 38 oyster growing estuaries in the Australian eastern state of NSW in 2008.

direct routine screening test for enteric viruses in shellfish food safety programs. Instead, reliance is still placed on tests for indicator bacteria either in water samples, oyster meat or both, as is the case in NSW.

This study reports on an epidemiological comparison of paired results of water and oyster meat indicator tests when conducted at multiple sites within three estuaries between 2000 and 2005 under the NSW Shellfish Program (NSW Food Authority 2001). We also used a subset of this data from selected oyster farming areas in the Port Stephens estuary to examine the statistical validity of an empirical model for relating these two fecal pollution indicators to environmental covariates.

MATERIALS AND METHODS

The selection of estuaries for this study was based on the following factors: the availability of a comprehensive data set that included parameters of interest and multiple sampling sites within regions of each estuary, geographic spread, and relative importance in state production. Three estuaries were selected:

the Hawkesbury River, Port Stephens and Manning estuaries situated on the central coast of NSW (see Fig. 1). Table 1 provides relevant fishery information for each of the study estuaries.

Port Stephens was divided into 6 zones: Karuah, Cromarty Bay, Tilligerry Creek, Swan Bay, North Arm Cove, and Tea Gardens as shown in Fig. 2.

The data used in this study were provided by the NSW Food Authority with permission from the Local Shellfish Programs (oyster farmers) for each estuary. Sampling conformed to the NSW Shellfish Program (SP). Each estuary was divided into sampling "zones" as specified in each estuary shellfish program. Samples were collected at the designated sentinel sampling sites, within or adjacent to oyster lease areas, within each "zone." Frequency of monitoring and water and oyster meat sampling zones varied but was at least fortnightly.

Live oyster samples were labeled and placed in a cold (0°C to 5°C) styrofoam box and delivered to the laboratory within 24 h of collection. A representative oyster sample was 12 average sized oysters, collected at random across each of the sampling sites in a sampling "zone." Samples were analyzed by an accredited laboratory using the Australian Standard method for examination for specific organisms —*E. coli* in bivalve molluscs—Rapid method (AS 1766.2.12). This method, known as the direct plating (DP) method (Anderson & Baird-Parker 1975), is an agar plate technique based on the production of indole, allowing the detection of *E. coli*. In brief, the oyster meat samples are blended and decimally diluted in peptone water (0.1%) then plated on a 0.45 mm membrane filter (MF) overlaying a tryptone bile agar. The plates are incubated at 44.5°C ± 0.5°C for 20–24 h. The MF is then removed and placed in indole reagent for 10–15 min and then dried. The pink/red colonies appearing on the MF are indole producers and are enumerated as *E. coli* Biotype I. Results are expressed as *E. coli* g^{-1} of fresh oyster meat. The method is faster than mean probable number (MPN) procedures (e.g., Andrews & Presnell 1972) and comparative evaluation found no significant differences in detection results Yoovidhya and Fleet (1981).

Water samples were collected in duplicate at each site according to the *NSW Shellfish Program Environmental Sampling for Shellfish Farming Course Handbook*. This guideline incorporates the Australian Standard ASNZF 5667.9—1998 Part 9 Guidance on Sampling from Marine Waters as well as other relevant information and guides. Samples were collected at designated sites by submerging a 250 mL sterile sample bottle to approximately elbow depth below the surface, whereas facing the bottle into the flow. All water samples were labeled and placed in a cold (0°C to 5°C) styrofoam box and delivered to the laboratory within 24 h of collection. Samples were analyzed for enumeration of fecal coliform units (thermotolerant coliforms and *E. coli*), measured as FCU $(100\ mL)^{-1}$, by an accredited laboratory using the Australian Standard membrane filtration method (AS 4276.7). A test portion of the sample is filtered through a membrane filter with a pore size 0.45 μm. The membrane is placed on Membrane Lauryl Sulphate agar (which contains lactose and an acid indicator) and incubated at 30°C for 4–6 h for resuscitation of the bacteria and then at 44°C to 44.5°C for 14–16 h. Colonies, corresponding to the viable organisms, grow on the filter and are counted. All flat lemon yellow colonies, 1–2 mm in diameter that grow on the filter, are counted as presumptive. Results are expressed as colony-forming units present in 100 mL of sample.

Salinity and temperature of the water sampling sites were measured by taking a 2-liter surface sample and readings taken immediately on board the boat with hydrometer and thermometer. The salinity was expressed as parts per thousand (ppt) and water temperature as degrees Celsius (°C).

Daily rainfall data was collated from the Australian Bureau of Meteorology for the Port Stephens catchment over the period during which the study was conducted. Three continuously recording weather stations are located within the catchment and the data used for rainfall were based on the closest available station, which in all cases was less than 10 km from the sampling site.

TABLE 1.

Selected study estuaries and fishery information.

Estuary	Estuary Water Area (ha)	Lease Area (ha)	Oyster Production (units/annum)
Manning	7,529	315	2 million
Hawkesbury	15,539	436	8 million
Port Stephens	57,231	650	15 million

Statistical Methods

There are several characteristics commonly used in epidemiological studies to describe the quality and usefulness of a diagnostic test. Accuracy can be expressed through sensitivity and specificity, positive and negative predictive values, and positive and negative diagnostic likelihood ratios. These are important population surveillance parameters that describe the test performance for a given reference population, having a given distribution of covariate factors, under defined conditions (Greiner & Gardner 2000, Bhopal 2002).

The water FCU test is the harvest management diagnostic test commonly recognized by the Australian and the United States of America shellfish programs. Provided there has been no "significant" recent rain or other potential pollution risks, this test is used to determine whether an area should be Open (–) or Closed (+) to harvest of oysters. Consequently the water FCU test is the nominative "reference test" in the SP.

The sensitivity of a test is formally defined as the proportion of true positives that yield a positive test in a contaminated/infected population. Conversely, specificity is the proportion of "true" negatives that actually test negative (or 1—false positive proportion). The likelihood ratio is a useful concept in epidemiology that represents the ratio of the probability of a "disease" being present to a disease being absent. In this study, the positive diagnostic likelihood ratio (DLR) represents the odds ratio that a positive test result will be observed in a Closed to harvest scenario *versus* the odds that the same result will be observed in an Open to harvest scenario. The negative DLR represents the odds ratio that a negative test result will be observed in a Closed to harvest scenario *versus* the odds that the same result will be observed in an Open to harvest scenario.

The calculation of these parameters using Table 2 is shown below:

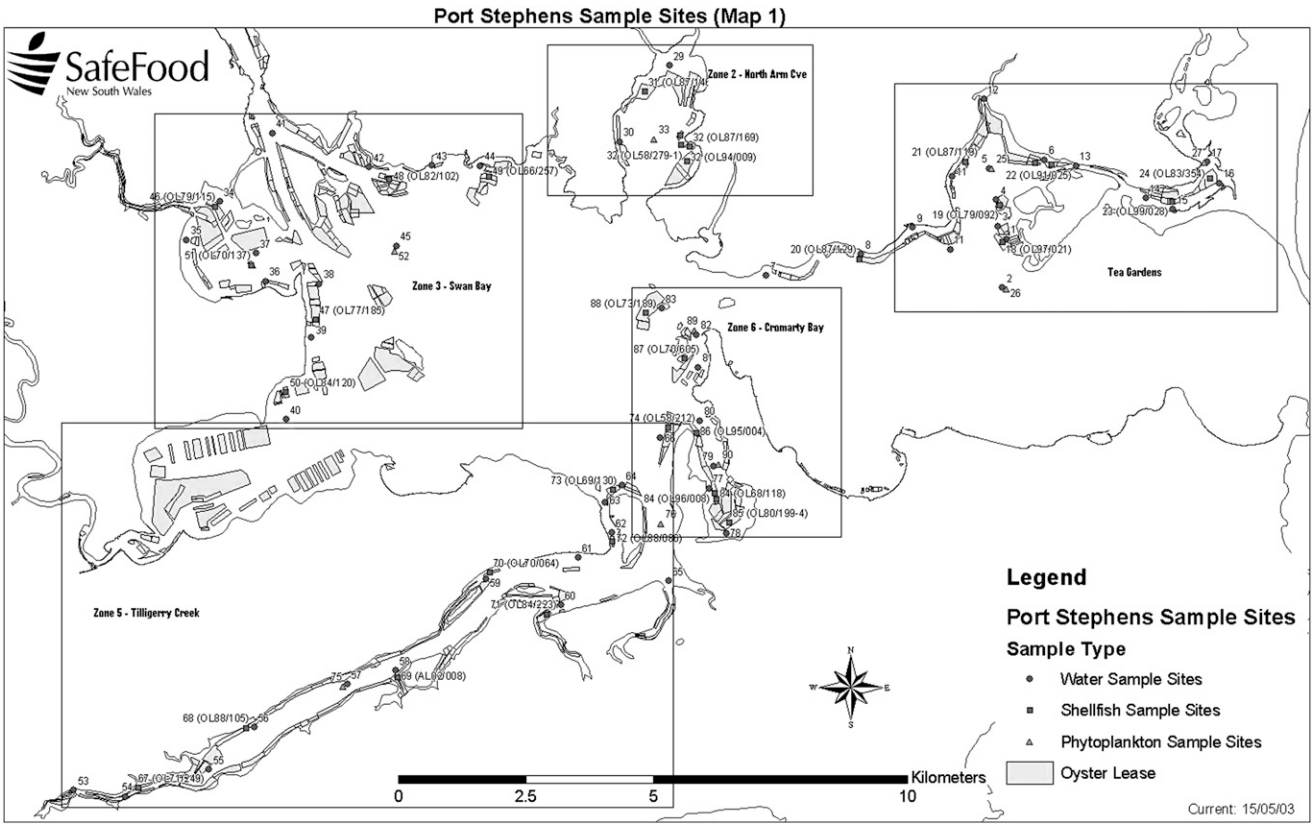

Figure 2. Zones and sampling sites in Port Stephens, NSW.

$$\text{Sensitivity} = \text{TP}/(\text{TP}+\text{FN})$$

$$\text{Specificity} = \text{TN}/(\text{TN}+\text{FP})$$

$$\text{Positive Diagnostic Likelihood Ratios} = (\text{TP}/[\text{TP}+\text{FN}]/\text{FP}/ \times [FP+TN]) = Sensitivity/(1-Specificity)$$

$$\text{Negative Diagnostic Likelihood Ratios} = (\text{FN}/[\text{TP}+\text{FN}])/ \times TN/[FP+TN]) = (1-Sensitivity)/Specificity.$$

The quality of oyster harvest areas are influenced by temporal changes estuary water quality caused by changes in catchment run-off. This is driven by rainfall, which in turn influences salinity of the estuary. Management tools that are used in the SP also include monitoring of rain events and salinity of the harvest area. The statistical validity of an empirical model for relating the two fecal pollution indicators (FCU and *E. coli*) to the environmental covariates, salinity, temperature, and rainfall was therefore also examined.

TABLE 2.

Table used for parameters that describe the quality and usefulness of a test.

		Reference Test Results	
		Closed(+)	**Open (–)**
New Test Results	**Closed(+)**	TP	FP
	Open(–)	FN	TN

TP = number of true positive specimens; FP = number of false positive specimens; FN = number of false negative specimens; TN = number of false negative specimens.

Port Stephens's estuary was selected for model development because of its large size and the presence of oyster farming lease areas within each of the four geomorphic zones that describe southeast Australian estuary systems (Roy et al. 2001). Five zones were chosen within the estuary to explore the degree of association between the indicator tests and temperature, rainfall, and salinity within the respective areas. Cross validation methods were used to select the best descriptive model to describe the relationship between the diagnostic tests and the environmental covariates (Efron & Tibshirani 1993). These methods compare different models based on their ability to predict responses using data independent of the data selected to develop the model. In particular, we use k-fold cross validation, in which the data set used to develop the model are first divided into a number of disjoint subsets, k, of data. Here, we set $k = 7$.

In brief, k-fold cross validation works as follows for a given model. Exclude from the data one of the subsets; fit the model to the retained subsets; use the fitted model to predict the response in the excluded subset and then repeat this for each subset. This then gives a predicted response for each observation within each

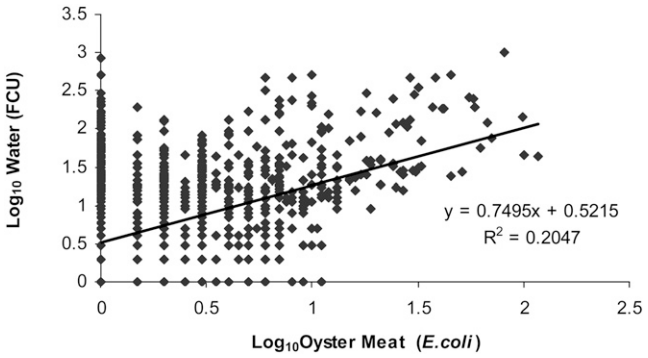

Figure 3. Comparison of paired samples (taken at same time and location) of Oyster *E. coli* and Water FCU tests (n = 1,875) showing low correlation of the two test methods.

subset. These are then used to determine the mean predicted square error (PSE) for the model, as well as a standard error for the mean PSE. This is then repeated for all the models in contention and the model with the minimum mean PSE is selected as the most likely model. We adopt the ''one standard error'' rule (Friedman et al. 2001), to select the most parsimonious model having mean PSE within one standard error of the model with the minimum mean PSE.

RESULTS

Correlation of Meat and Water Samples

A total of 1,875 paired samples for FCU and *E. coli*, spanning the five years of sample collection in three estuaries, were examined. This data set were used for correlation analysis. Ideally, the fecal pollution indicators (FCU in water and *E. coli* in meat) used in the SP should correlate closely for the indicators to be considered meaningful as either or both can be used for opening/closing estuaries for harvest in NSW. Correlation analysis of the two indicators was performed (see Fig. 3) using 1,875 ''paired'' oyster and water samples. A very low correlation ($R^2 = 0.20$) was found between the logs (to the base 10) of the two indicators. It can also be noted that there are a large number of readings on both axes suggesting that equivalency in positive results for both tests is low.

Agreement between the Two Indicator Tests

In the Hawkesbury River and Port Stephens data sets daily rainfall data in the catchment were recorded. We used this to block the data sets. Samples collected where no rainfall had occurred in the previous 72 h and when the total rain in the previous week was <10 mm were deemed to be ''Open'' to harvest. The remainder were deemed to be collected when ''Closed'' to harvest. This blocking was used to reduce ''likely prevalence'' of positive tests.

In accordance with the SP, a positive result for an oyster *E. coli* test of >2.3 *E. coli*·g^{-1} of oyster meat is determined as the cut-off value. Oyster meat values less than and greater than this *E. coli* cut-off were classed as ''low'' and ''high'' respectively. Similarly in accordance with the SP, the Water FCU cut-off value for a positive test is >14 FCU·100 mL^{-1}. Water values less than and greater than this FCU cut-off were classed as ''low'' and ''high'' respectively.

Table 3 presents the results of the two tests in the different estuaries. Overall, a total of $n = 848$ samples were compared. Seventy-three percent of the results gave matching negative readings for both tests. However, over half of the water FCU samples tested positive when the oyster meat *E. coli* test was negative. In some zones, the discordant diagonal results were even greater, indicating substantial disagreement between the tests. For example in Port Stephens 10 out of 12 oyster *E. coli* tests were ''high'' when no rain had occurred in the region for more than a week but the corresponding water FCU readings were all below the regulatory cut-off point. Conversely, 44 water FCU tests were high in the Hawkesbury River when no rain had occurred in the region for more than a week but corresponding oyster meat *E. coli* readings were below the regulatory cut-off point.

Comparison of Kappa Values between the Two Indicator Tests

Comparison of the two diagnostic tests was also measured by the quotient called kappa (κ). Kappa scores are a measure of the agreement beyond chance divided by the maximum chance agreement (Cohen 1960). κ-scores have been calculated

TABLE 3.

Comparative analysis of corresponding meat *E. coli* and FCU water samples for estuaries.

Hawkesbury "Open"	**FCU High**	**FCU Low**	**Tilligerry**	**FCU High**	**FCU Low**
Oyster *E. coli* high	13	11	Oyster *E. coli* high	12	20
Oyster *E. coli* low	44	306	Oyster *E. coli* low	13	56
Hawkesbury "Closed"	**FCU High**	**FCU Low**	**Tea Gardens**	**FCU High**	**FCU Low**
Oyster *E. coli* high	15	2	Oyster *E. coli* high	0	4
Oyster *E. coli* low	13	11	Oyster *E. coli* low	4	15
Cromarty	**FCU High**	**FCU Low**	**Port Step. "Open"**	**FCU High**	**FCU Low**
Oyster *E. coli* high	4	7	Oyster *E. coli* high	2	10
Oyster *E. coli* low	3	62	Oyster *E. coli* low	8	197
Manning	**FCU High**	**FCU Low**	**Total**	**FCU High**	**FCU Low**
Oyster *E. coli* high	18	20	Oyster *E. coli* high	66	59
Oyster *E. coli* low	13	134	Oyster *E. coli* low	98	625

TABLE 4.

Computation of kappa as a measure of agreement between the two diagnostic tests (Oyster *E. coli* and Water FCU) in different data sets.

Location	κ	Location	κ
Hawkesbury "Open"	0.25	Tilligerry	0.20
Hawkesbury "Closed"	0.31	Pindimar	0.21
Cromarty	0.37	Corrie Island	0.13
Manning	0.41	Port Stephens "Open"	0.14
		Total	**0.34**

separately for each of the 8 data sets and as a combined total score and are presented in Table 4.

Agreement between the two tests varied among data sets ($0.14 < \kappa < 0.41$) and all were less than 0.5. The κ-score for the combined 8 data sets ($n = 1{,}065$) was 0.34 indicating that the intertest agreement is low.

Evaluating Equivalency of the E. coli Test

Indirect measures to determine the apparent sensitivity and specificity of a diagnostic test require a relatively large number of test results ($n > 1{,}000$) with an assumed prevalence of "contamination" less than 1%. It also requires that the test has high sensitivity (Martin et al. 1987). To satisfy these requirements, results of paired tests were selected from the Hawkesbury River and Port Stephens data sets for the period from May 20, 2003 to March 31, 2005 when no rainfall had occurred in the previous 72 h and total rain in the previous week was <10 mm. Because the Water FCU test is the management tool used in the Australian and United States programs for the approval of harvest, we use that as the "reference" test for comparing the oyster E. *coli* test. Results are presented in Table 5.

The sensitivity of the *E. coli* test was calculated using this data set to determine the probability of a positive *E. coli* result when a positive FCU test would prevent oysters being harvested from the area. The specificity of the *E. coli* test was calculated to determine the probability of a negative *E. coli* test when a negative FCU test would allow harvest. In addition the positive and negative DLR was calculated to determine respectively:

The odds ratio of a positive *E. coli* test when FCU test is positive versus a positive *E. coli* test when the FCU test is negative;

TABLE 5.

Data used to evaluate the diagnostic tests for samples taken when no rain had occurred in the previous 72 Hours and < 10mm total rain in the previous week.

Low Prevalence Samples	Water FCU High	Water FCU Low	Total
	Hawkesbury		
Oyster *E. coli* high	22	100	122
Oyster *E. coli* low	33	1004	1037
Total - Hawkesbury	**55**	**1104**	**1159**
	Port Stephens		
Oyster *E. coli* high	2	10	12
Oyster *E. coli* low	8	197	205
Total – Port Stephens	**10**	**207**	**217**

TABLE 6.

Equivalence measures and associated standard errors for *E. coli* tests in Hawkesbury and Port Stephens estuary respectively.

Performance Measure	Hawkesbury	Port Stephens
Sensitivity	40%±7%	20%±3%
Specificity	91%±2%	95%±3%
Positive DLR	4.4	4.0
Negative DLR	0.67	0.85

The odds ratio of a negative *E. coli* test when FCU test is positive versus a negative E. *coli* test when the FCU test is negative.

The results for these measures are presented in Table 6.

The oyster *E. coli* test has a low apparent sensitivity (40% in Hawkesbury and 20% in Port Stephens) and a reasonable specificity (91% in H. and 95% in PS). The low sensitivity indicates that the probability is high of a negative oyster *E. coli* test when a water FCU test is positive and the specificity indicates there is smaller probability of a positive oyster *E. coli* test when water FCU test is negative. However, the positive DLR for Port Stephens and Hawkesbury indicate that for every 5 oyster *E. coli* tests conducted when the harvest area is deemed open by the FCU water test, approximately 1 oyster meat test will test positive. Conversely, the negative DLR indicate that for every for every 4 oyster *E. coli* tests conducted when the harvest area is deemed closed by the FCU water test approximately 1 will test negative. In epidemiological applications a good diagnostic test ideally has a positive DLR greater than 10 and a negative DLR less than 0.1 (Hodgson & Tannock 2005).

Relationship between Diagnostic Tests and Environmental Covariates

Samples were collected under the auspices of the NSW Shellfish Program from designated sites within or adjacent to oyster lease areas assigned to "zones" specified in the Port Stephens shellfish program (see Table 7). Data was sourced from the NSW Food Authority and local shellfish programs established under the Program.

The maximal model we consider is:

$$Y = \text{mean} + \text{Zone} + \text{Zone} * [\text{Temp} + \text{Sal} + \text{RWeek} + \text{R0} + \text{R24} + \text{R72} + \text{JDay} + \text{Cos(SDay)} + \text{Sin(SDay)}] + \textbf{\textit{Zone: Site}} + \textbf{\textit{FDay}} + \textbf{\textit{Zone: FDay}} + \textit{error}$$

In brief, this full model allows the response Y, equal to $\log_{10}$ (FCU +1), within each of the five zones (zones 2, 3, 5, 6, and 7)

TABLE 7.

Zone sampling statistics used in development of the predictive model.

Zone	Geographic Name	Number Sampling days	Number sites Sampled—FCU
2	North Arm Cove	14	3
3	Swan Bay	25	12
5	Tilligerry	30	14
6	Cromarty Bay	13	8
7	Karuah River	15	5

to have a baseline response for each Zone (mean + Zone) and to be linearly related to temperature (Temp), salinity (Sal), previous week's rainfall (RWeek), previous day's rainfall (R0), rainfall 24–48 h prior to measurement (R24) and rainfall 72–168 h prior to measurement (R72). Also, a linear trend in time (JDay) and a seasonal trend with a yearly cycle (Cos(SDay) + Sin(SDay)) is included. The inclusion of "Zone *" in the earlier shown model indicates that the dependence on the above terms is possibly different within each zone. The remaining terms in the model, those in bold italics, are included as random effects, which are sources of variation. ***Zones:Site*** allows for variation across sites within each zone, ***FDay*** is a factor associated with each unique sampling day, ***Zones:FDay*** allows for extra variation across days within each zone and finally ***error*** is the random error in the model. All random terms are assumed independent, which is a first approximation given that there could be some auto-correlation with FDay effects.

Treating Cos(SDay) + Sin(SDay) in the earlier illustrated model as a single term, because both are required to model the seasonal trend, we have that the above model contains forty (40) fixed effects terms (excluding mean + Zone) that can either be retained or omitted. Hence we have 2^{40} possible models if all are to be compared. This is computationally impossible and hence we use a step-down k-fold cross validation procedure to select a suitable model. That is, we begin with the maximal model earlier and estimate its mean PSE and associated standard error. Then each submodel of this maximal model containing all but one of the fixed effects terms is fitted and the mean PSE and associated standard error are estimated. The model from this subset of models with minimum mean PSE is then selected as the "maximal-model" for the next stage. This process is then repeated until the mean PSE for all subsequent submodels begin to become unacceptably large.

The results of applying this step-down k-fold cross validation procedure ($k = 7$) up to the stage where omitting further terms leads to an unacceptable increase in mean PSE from the full model we are left with the following reduced model in which significantly selected coefficients for each parameter in each Zone are identified as **at(Zone…)**.

$$\begin{aligned}\mathrm{Y} = {} & \mathrm{Zone} + \mathrm{at}(\mathrm{Zone}, \mathrm{c}[2,3,5,6]) : \mathrm{Sal} + \mathrm{at}(\mathrm{Zone}, 7) : \mathrm{Temp} \\ & + \mathrm{at}(\mathrm{Zone}, \mathrm{c}[5,6,7]) : \mathrm{R0} + + \mathrm{at}(\mathrm{Zone}, \mathrm{c}[5,6]) : \mathrm{R24} \\ & + \mathrm{at}(\mathrm{Zone}, \mathrm{c}[3,5]) : \mathrm{R48} + \mathrm{at}(\mathrm{Zone}, \mathrm{c}[2,3]) : \mathrm{R72} + \mathrm{at}(\mathrm{Zone}, 3) : \\ & \mathrm{JDay} + \mathit{Zone:Site} + \mathit{FDay} + \mathit{Zone:FDay} + \mathit{error}\end{aligned}$$

Fitting this model to the full data set we have a predictive model with the coefficients found in Table 8 with the Z-Ratio being the value of the Coefficient divided by the Standard Error. This ratio provides an absolute measure of the robustness of each variable in the model.

Here the notation is, for example, that salinity is a useful predictor for FCU at zones 2, 3, 5, and 6 whereas not so at zone 7, after adjusting for the other variables in the model. At Zone 7 Temp seems a better predictor than salinity after adjusting for rainfall. At Zone 3 we see that there is a trend with time (JDay effect). Estimates of the variance parameters in the model are given in Table 9.

From the earlier shown model we see that YFCU (hence FCU) decreases with increasing salinity at Zones 2, 3, 5, and 6 and with increasing temperature at Zone 7 when the rainfall variables are fixed (and the trend in time at Zone 3 is removed).

TABLE 8.

Selected variables and respective coefficients derived from the model.

Variable	*Co-efficient*	*Std. Error*	*Z Ratio*
at(Zone3):JDay	–0.0010	0.0005	–2.2
at(Zone3):R72	0.0117	0.0029	4.0
at(Zone2):R72	0.0228	0.0059	3.9
at(Zone5):R48	0.0155	0.0066	2.4
at(Zone3):R48	0.0463	0.0066	7.0
at(Zone6):R24	0.0405	0.0101	4.1
at(Zone5):R24	0.0363	0.0073	4.9
at(Zone7):R0	0.0528	0.0119	4.4
at(Zone6):R0	0.0185	0.0077	2.4
at(Zone5):R0	0.0262	0.0065	4.0
at(Zone7):Temp	–0.1653	0.0208	–7.9
at(Zone6):Sal	–0.1604	0.0226	–7.1
at(Zone5):Sal	–0.1221	0.0094	–13.1
at(Zone3):Sal	–0.1368	0.0112	–12.2
at(Zone2):Sal	–0.1702	0.0220	–7.7
Zone2	5.8192	0.7692	7.6
Zone3	5.2955	0.4203	12.6
Zone5	5.3132	0.3134	17.0
Zone6	5.5391	0.7493	7.4
Zone7	5.5200	0.5341	10.4

From the model we have, for example, that the predictive model for Zone 2, is

$$\text{Predicted Y at Zone 2} = 5.8192 - 0.1702\,\mathrm{Sal} + 0.0228\,\mathrm{R72}$$

Based on the model we can see how the observed values compare with the fitted values at each Zone. This is illustrated in Figure 4.

DISCUSSION

Classification decisions for oyster growing and harvest areas in a shellfish program are contingent on a number of critical factors. These include, but are not limited to, actual and potential pollution sources, and their quantification, as well as meteorological and seasonal factors affecting the harvest area in conjunction with periodic monitoring of growing areas by microbiological and hydrochemical assays of the waters (Dressel & Snyder 1991). This multidisciplinary risk reduction approach has become the accepted strategic philosophy in recognition of the inadequacy of a classification system based merely on results of bacteriological fecal pollution indicators measured in the water or the bivalves (Ferguson et al. 1996, Henshilwood et al. 1998, Guyader et al. 1998, Dore et al. 2000, Rodgers 2001).

TABLE 9.

Estimates of the variance parameters in the model.

Source of Variance	*Variance Estimate*	*Std. Error*	*Z-Ratio*
Zone:Site	0.2427	0.0601	4.0
FDay	0.3324	0.0791	4.2
Zone:FDay	0.2891	0.0591	4.9
R!variance	0.6204	0.0202	30.7

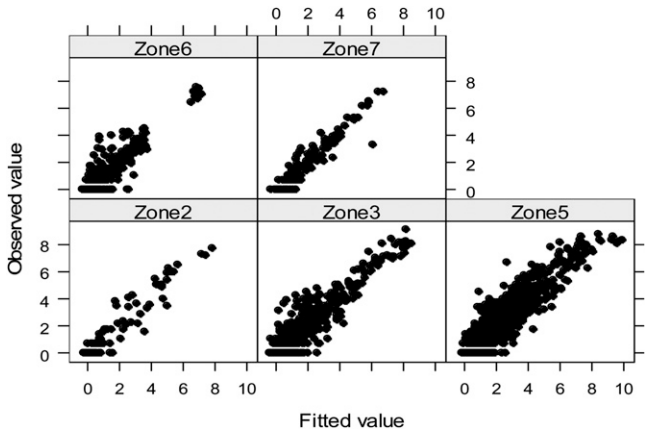

Figure 4. Comparison of observed and fitted values of Log_{10}(FCU + 1) at each Zone.

However, a major problem with present water pollution warning systems in shellfish programs is that bacterial indicator concentrations are spatio-temporally variable and most sampling is too infrequent to transcend any spikiness. To overcome this constraint, shellfish programs use predictive models to establish management plans based on climatic and biophysical triggers for particular shellfish areas to minimize public health risk. The science of epidemiology is useful in the development and understanding of these predictive models as it seeks to describe, understand and use patterns of ''disease'' to improve health. Epidemiology is in many respects an ecological discipline with the unit of analysis always the group (Bhopal 2002). A good epidemiological exposure variable reflects the purposes of epidemiology. It is measurable accurate, differentiates populations in their experience of disease or health and generates testable aetiological hypotheses, or helps in developing management plans to prevent and control disease (Bhopal 2002).

In this context, we have examined here the two fecal pollution indicators currently used in the NSW Shellfish Program and found that:

1. Both graphical representation of pairs of observations of the two tests, where a ''pair'' is an oyster flesh *E. coli* count and FCU count on water sample from nearby sampled at the same time, and formal estimation of the correlation coefficient ($r^2 = 0.21$ and $n = 1{,}875$) indicate that the two tests are poorly related. This poor agreement between the two tests is reflected in the low range of κ-scores $0.14 < \kappa < 0.41$determined in three estuaries in the NSW central coast ($n_{total} = 1{,}065$).
2. Calculation of sensitivity and specificity for the *E. coli* test gave a low sensitivity estimate indicating that the probability of an oyster *E. coli* test being negative when a water FCU test is positive is high.
3. Calculation of the diagnostic likelihood ratios revealed that for every 4 oyster *E. coli* tests conducted when the harvest area is deemed closed by the FCU water test, approximately 1 will test negative. Conversely, for every 5 oyster *E. coli* tests conducted when the harvest area is deemed open by the FCU water test, approximately 1 will test positive. A substantial unexplained systematic error is occurring in the test. This results in disruption to trade including closures and product recall.
4. The FCU water test is highly correlated with rainfall and salinity as would be expected. In certain cases seasonal trends (indicated by water temperature) are also apparent. Conversely, the oyster *E. coli* test is poorly correlated with these environmental covariates.

There could be several reasons for the poor correspondence between these tests. In periods of high rainfall when FCUs are elevated, estuarine waters are fresh and physiologically hostile to oysters. Sydney rock oysters are remarkable in that they are capable of remaining closed for up to three weeks if conditions are suboptimal (Oyster Culture Commission 1877). As well, positive oyster *E. coli* could come from bird fecal materials as birds tend to perch on oyster infrastructure. Finally the sampling, handling, transport, and analysis of oyster meat samples seem more problematic than those involved in water samples.

CONCLUSION

By testing a small subset of the total ''population,'' oyster meat and water diagnostic tests make an inference about all oysters and the water in the locality (zone) where the testing is done. In veterinary epidemiological terms this is known as ''herd testing.'' In the water test, an integrative environmental indicator is used to infer the quality of the harvest area; in the oyster test a direct inference is made on the oyster population as a whole by subsampling. The source of sampling error at the level of the ''oyster lease'' and ''water sampling zone'' is part of the systematic error of the ''herd test.''

The results of the analyses presented here indicate that the current monitoring program using meat testing has considerable systematic error that has not been identified in the NSW Shellfish Program. The comparative performance of the environmental test for Water FCU with the oyster *E. coli* test suggests that there are significant performance differences between the tests. The fact that the FCU water tests were correlated with environmental parameters, local rainfall, salinity, and water temperature as expected, whereas the oyster meat *E. coli* tests were not, suggests the former has some credibility. The oyster meat test tacitly assumes that the oysters retain *E. coli* following high runoff events when the water column may have cleared. However, the results demonstrate that even during extended low rainfall periods, the probability of getting a positive oyster meat *E. coli* test when the water FCU test was negative was at least one in five meat tests and seemingly random in occurrence.

Sampling, handling, and laboratory errors as well as possible confounding factors such as contamination from birds should be examined for contribution to systematic errors in this test. Oyster condition, presence/absence of mudworm, and bird roosts on leases are possible confounders that could be investigated.

The focus of the shellfish program is on the quality/suitability of the harvest area. The results of these analyses suggest that the Water FCU test, as adopted by the United States and Australian Shellfish Programs provides a more reliable measure of exposure of oysters to fecal pollution than the combined dual testing of oyster meat and water adopted by the NSW Shellfish program.

The routine dual use in harvest management in the NSW SP results in routine uncertainty that creates significant unpredictability in trade and major costs. In turn, this uncertainty and

cost result in levels of commercial risk that could be unsustainable. The study highlights the need for the performance of tests to be appraised in shellfish quality assurance programs. In so doing it may be possible to maintain public health standards while minimizing unnecessary and costly disruptions in the trade of fresh oysters.

ACKNOWLEDGMENTS

The authors acknowledge Remy van de Ven and David Jordan of NSW Department of Primary Industries for their assistance in the statistical analysis and comments on an earlier related paper prepared for the Eleventh International Symposium on Veterinary Epidemiology and Economics held in Cairns Australia in 2006.The authors also acknowledge the NSW Food Authority for the data provided by them for use in this paper. Thanks to NSW oyster farmers too numerous to mention for support and permission to use their results and a special thanks to Daniel Ogburn and Ben Ogburn who assisted in the data entry and a lifetime of patient visits to NSW oyster farms.

LITERATURE CITED

Anderson, J. M. & A. C. Baird-Parker. 1975. A rapid and direct plate method for enumerating *Escherichia coli* biotype I in food. *J. Appl. Bacteriol.* 39:111–117.

Andrews, W. H. & M. W. Presnell. 1972. Rapid Recovery of *Escherichia coli* from Estuarine Water. *Appl. Environ. Microbiol.* 23:521–523.

Bhopal, R. 2002. Concepts of epidemiology: an integrated introduction to the ideas, theories, principles and methods of epidemiology. New York: Oxford University Press.

Cohen, J. 1960. A coefficient for agreement of nominal scales. *Educ. Psychol. Meas.* 20:37–46.

Dorairaj, S. & B. M. Miller. 2001. Major issues with human enteric virus behaviour within estuarine environments. Technical Report 2000/02. The University of New South Wales Water Research Laboratory. Sydney, Australia.

Dore, W. J., K. Henshilwood & D. N. Lees. 2000. Evaluation of F-specific RNA bacteriophage as a candidate human enteric virus indicator for bivalve molluscan shellfish. *Appl. Environ. Microbiol.* 66:1280–1285.

Dressel, D. M. & M. I. Snyder. 1991. Depuration—the regulatory perspective. In: W. S. Otwell, G. E. Rodrick & R. E. Martin, editors. Molluscan shellfish depuration. Boca Raton, Florida: CRC Press. pp. 19–23.

Efron, B. & R.J. Tibshirani. 1993. An introduction to the bootstrap. New York: Chapman and Hall.

Ferguson, C. M., B. G Coote, N. J. Ashbolt & I. M. Stevenson. 1996. Relationships between indicators, pathogens and water quality in an estuarine system. *Water Research* 30:2045-2054.

Friedman, J.H., T. Hastie & R.J. Tibshirani. 2001. The Elements of Statistical Learning. New York: Springer.

Green, D. H. & G. D. Lewis. 1999. Comparative detection of enteric viruses in wastewaters, sediments and oysters by reverse transcription-PCR and cell culture. *Water Res.* 33:1195–1200.

Greiner, M. & I. A. Gardner. 2000. Epidemiological issues in the validation of veterinary diagnostic tests. *Prev. Vet. Med.* 45:3–22.

Griffin, D. W., E. K. Lipp, K. McLaughlin, M. R. Rose & J. B. Rose. 2001. Marine recreation and public health microbiology: quest for the ideal indicator. *Bioscience* 51:817–825.

Guyader, F., L. Haugarreau, L. Miossec, E. Dubois & M. Pommepuy. 2000. Three-year study to assess human enteric viruses in shellfish. *Applied and Environmental Microbiology* 66:3241–3248.

Guyader, F., J. Miossec, L. Haugarreau, E. Dubois, H. Kopecka & M. Pommepuy. 1998. RT-PCR evaluation of viral contamination in five shellfish beds over a 21-month period. *Water Sci. Technol.* 38:45–50.

Hallegraeff, G. M. 2002. Aquaculturists' guide to harmful Australian microalgae. School of Plant Science, University of Tasmania, Hobart, Australia. 136 pp.

Henshilwood, K., J. Green & D. N. Lees. 1998. Monitoring the marine environment for small round structured viruses (SRSVs): A new approach to combating the transmission of these viruses by molluscan shellfish. *Water Sci. Technol.* 38:51–56.

Hodgson, D. C. & I. F. Tannock. 2005. Guide to studies of diagnostic tests, prognostic factors, and treatments. In: I. Tannock, R.P. Hill, R.G. Bristow & L. Harrington, editors. The Basic science of oncology, 4th ed. New York: McGraw-Hill. pp. 489–508.

Jackson, K. L. & D. M. Ogburn. 1999. Review of depuration and its role in shellfish quality assurance. NSW Fisheries Final Report Series No. 13. FRDC Project No. 96/355, 77 pp.

Jiang, S. C., R. Nobel & W. Chu. 2001. Human adenoviruses and coliphage in urban runoff-impacted coastal waters of southern California. *Appl. Environ. Microbiol.* 67:179–184.

Lees, D. N. 2000. Viruses and bivalve shellfish. *Int. J. Food Microbiol.* 59:81–116.

Martin, S. W., A. H. Meek & P. Willeberg. 1987. Veterinary epidemiology—Principles and methods. Iowa State University Press, Ames, Iowa. 343 pp.

New South Wales Food Authority. 2001. NSW Shellfish Program Operations Manual. Accessible at: http://www.foodauthority.nsw.gov.au/industry/fb-shellfish.asp.

Noble, R. T. & J. A. Fuhrman. 2001. Enteroviruses detected by reverse transcriptase polymerase chain reaction from the coastal waters of Santa Monica Bay, California: low correlation bacterial indicator levels. *Hydrobiol.* 460:175–184.

Ogburn, D. M., I. White & D. McPhee. 2007. The disappearance of Oyster Reefs from Eastern Australian Estuaries – Impact of Colonial Settlement or Mudworm Invasion? *Coast. Manage.* 35:271–287.

Oyster Culture Commission. 1877. Report of the Royal Commission, appointed on September 29, 1876 to inquire into the best mode of cultivating the oyster together with the minutes of evidence, and appendices. Charles Potter, Acting Govt. Pr., Sydney, NSW. 73 pp.

Rodgers, C.J. 2001. The NSW Shellfish quality assurance program: an operational review. Final report January 2001. NSW safefood production, Sydney.

Roy, P. S., R. J. Williams, A. R. Jones, I. Yassini, P. J. Gibbs, B. Coates, R. J. West, P. R. Scanes, J. P. Hudson & S. Nichol. 2001. Structure and function of South-east Australian estuaries. *Estuar. Coast. Shelf Sci.* 53:351–384.

Wetz, J. J., E. K. Lipp, D. W. Griffin, J. W. Lukasik, D. Wait, M. D. Sobsey, T. M. Scott & J. B. Rose. 2004. Presence, infectivity, and stability of enteric viruses in seawater: relationship to marine water quality in the Florida Keys. *Mar. Pollut. Bull.* 48:698–704.

Yoovidhya, T. & G. H. Fleet. 1981. An evaluation of the A-1 most probable number and the Anderson and Baird-Parker plate count methods for enumerating *Escherichia coli* in the Sydney rock oyster (*Crassostrea commercialis*). *J. Appl. Bacteriol.* 50:519–528.