Evaluation of new, rapid microbiological methods for measuring recreational water quality

John F. Griffith, Stephen B. Weisberg and Charles D. McGee[1]

ABSTRACT - Public health officials routinely measure fecal indicator bacteria to assess beach water quality, but present laboratory methods require about 24 hours to obtain a result, delaying issuance of health warnings until the day following sampling. New, more rapid measurement methods that would allow for same day warnings are under development, but have not yet undergone independent testing. Here we evaluate four of these new methods: immunomagnetic separation coupled with ATP bioluminescence, flow cytometry, quantitative polymerase chain reaction (Q-PCR) and dual-wavelength fluorimetry (DWF). Testing was conducted in two phases. In the first phase, developers of each new method processed 54 blind samples. Results were then compared to those produced by five laboratories using two traditional methods of analysis. Samples included both natural and laboratory-created samples, ensuring method evaluation over a range of concentrations, matrices and interferences. None of the new methods performed well enough to replace existing methods at present, but there were encouraging results for two methods. Q-PCR produced results within the range of the reference labs for two-thirds of the samples, but overestimated others, particularly those with complex matrices. DWF had the best precision among the new methods and was more precise than some of the laboratories using traditional methods, but also produced results that were generally higher than existing methods, especially for samples that contained urban runoff. In the second phase of testing, practitioners of traditional culture-based methods were trained in the new methods and asked to process nine blind samples. These practitioners produced results comparable to those produced by the method developers. While there is need for performance improvement in each of the new methods, technology transfer does not seem to be an important impediment to their adoption.

INTRODUCTION

California's public health officials routinely measure fecal indicator bacteria to assess recreational water quality. State Health Department regulations require weekly measurements of indicator bacteria (total coliforms, fecal coliforms, and enterococci) at high-use beaches. These regulations further require that the public be warned of possible health risks if any of these bacterial indicators exceed threshold values that were established through epidemiological studies.

California's investment in these water quality-monitoring programs could be improved if the laboratory time for enumerating bacteria was shortened. United States Environmental Protection Agency approved methods for measuring indicator bacteria in recreational waters require an 18- to 96-h incubation period, while several recent studies have shown that temporal changes in indicator bacteria levels in beach water occur much more rapidly (Leecaster and Weisberg 2001, Boehm *et al.* 2002). Thus, contaminated beaches can be open during the incubation period and be clean by the time warnings are posted. This time lag also inhibits tracking of contamination sources, since the signal can dissipate before upstream tracking is initiated. Lacking a more rapid method, investigators are unable to follow the trail of contamination back to its origin.

Technological advances provide new opportunities for measuring bacteria more rapidly. Whereas present methods rely on culturing bacteria and measuring growth or metabolic activity, new methods allow direct measurement of cellular attributes such as genetic material or surface immunological properties. Removing the extended incubation step allows these methods potentially to provide results in less

[1]*Orange County Sanitation District, 8127 Ellis Ave., Fountain Valley, CA 92708*

than 4 h, with sufficient time for managers to take action to protect the public health (i.e., post warning signs or close a beach) on the same day that water samples are collected.

Genetic and immunological methods have advanced considerably for use in disciplines such as the food service and hospital industries (Fung 2002, NRC 2004). Effort has recently begun to focus on application of these new methods for recreational water quality testing (Noble and Weisberg in press). Water testing presents challenges not frequently encountered in other fields, such as complex sample matrices and the presence of other potentially confounding native bacterial species. As such, extensive testing of these methods is needed to ensure that they provide comparable reliability to the culture-based methods they are intended to replace. Here, we present a study evaluating four new rapid methods to assess whether they are suitable replacements for existing methods.

METHODS

Four new rapid methods were tested. The first was immunomagnetic separation coupled with ATP Bioluminescence (IMS/ATP). The IMS/ATP method uses magnetic beads coated with antibodies specific to *Enterococcus*. to bind and capture the target bacteria in a water sample. Once bound, the magnetic bead/antibody/bacteria complexes are pulled out of solution using a powerful magnet. Now separated from the rest of the bacterial population, the captured cells are enzymatically-lysed, releasing their ATP into solution. This ATP is then quantified via a bioluminescent assay. The amount of ATP is converted to the number of bacteria captured following a calibration curve established using stock cultures of known concentration.

The second was a flow cytometry (FC) method, which employed the Advanced Analytical Technologies (Ames, IA) RBD 3000 instrument. Like IMS/ATP, this method derives its specificity from antibodies specific to *Enterococcus*. Once captured by the first antibody, the target cells are labeled using a secondary antibody containing a fluorogenic tag that, when excited by a laser in the instrument, allows the cells to be enumerated in the flow cytometer.

The third method was quantitative polymerase chain reaction (Q-PCR), which detects and enumerates unique genetic targets found in *Enterococcus*. The bacteria are first captured on a filter. The filter containing the bacteria is then subjected to bead beating, which mechanically lyses the cells and releases their DNA into solution. This DNA then is used in the quantification step, where enterococcal DNA is simultaneously amplified and measured using the Taqman® system of fluorescent probes (Applied Biosystems, Foster City, CA) and the advanced optics of the Q-PCR instrument (Cepheid, Sunnyvale, CA).

The last method was dual-wavelength fluorimetry (DWF), which relies on the same sugar-fluorophore substrate as is used in the commercially available IDEXX™ assays, but advances the detection process. Through use of a novel fluorometer, the method simultaneously measures the rate at which bacteria take up the chromogenic substrate as well as the rate at which the fluorescent byproduct of substrate metabolism appears. This ratiometric measurement allows detection and enumeration of target bacteria in a matter of only four hours.

Evaluation Approach

The evaluation approach was to assess equivalency with existing methods through simultaneous processing of water samples using both new and existing methods of enumerating fecal indicator bacteria. Samples processed included both natural samples and laboratory-created samples, to ensure that a range of conditions was evaluated. Laboratory-created samples were included because they offer the ability to control the number of indicator organisms and potentially interfering contaminants present, but they do not completely mimic natural conditions. Environmental water samples were included because they contain complex combinations of interferences that cannot be duplicated in artificial samples, though they offer less control over specific variables that need to be evaluated.

Testing occurred in two phases. The first phase involved application of the new methods by the experts that developed them. The second phase involved application of the methods by senior members of several local microbiology laboratories, who would become likely users of these methods should they be approved. The goal of the second phase was to assess whether these new technologies are readily transferable to local practitioners.

In phase one, participants analyzed 54 blind samples consisting of triplicates of each of 18 different test samples. Sample processing occurred over 3 d, with triplicates of each of six samples processed

on each day. Processing occurred over 3 d because participants identified that 18 samples were the most they could analyze within the 4-h time frame without duplicative equipment and personnel.

Six of the eighteen Phase I samples consisted of laboratory-created samples in a seawater matrix inoculated with differing levels of fecal contamination. Seawater used to prepare these samples was collected from 5 km offshore in an area known to be free from allochthonous fecal contamination. Three of these samples were inoculated with treated wastewater effluent from the Orange County Sanitation District (OCSD) in Fountain Valley, California. Three were inoculated with urban runoff collected from the Seventh Street drain, flowing into the Los Angeles River.

Four additional samples were a seawater matrix inoculated with sewage effluent, but with a more complex matrix. Two of these were created using the offshore seawater and OCSD effluent, but mixed humic acids (Sigma, St. Louis, MO) added at a concentration of 0.001% w/v. The other two were prepared using seawater collected at a nearshore location known to have historically high levels of suspended solids, but low levels of fecal bacteria. The inoculum for these samples was effluent from the Los Angeles County Sanitation District's Carson wastewater treatment plant.

Four samples were natural samples, including a shoreline sample collected at Doheny State Beach in Dana Point, California, and urban runoff samples from three freshwater locations: Yorktown Drain in Huntington Beach, California; Santa Ana River in Fountain Valley, California; and San Juan Creek in Dana Point, California.

The last four samples were different types of blanks. These consisted of sterile phosphate-buffered saline (PBS); uninoculated offshore seawater, 0.2 μm filtered offshore seawater, and 0.2 μm filtered urban runoff from Yorktown Drain.

Nine samples were processed in Phase II, in which local practitioners implemented the new methods. These consisted of triplicates of each of three different test samples in a seawater matrix. These samples were prepared as in Phase I, inoculating OCSD effluent into offshore seawater. All samples were blinded.

Phase I testing took place on June 2, 3, and 4, 2004. Samples were created or collected between 6:00 a.m. and 9:00 a.m. each day and distributed by 11:00 a.m. to each of the five local labs that used routine methods. All participants began processing samples at the same time and processed samples in numbered order to minimize any concentration differences that might have developed from degradation during sample transport or laboratory holding.

Phase II testing was conducted on June 9, 2004. Local users, each a senior member of a large southern California microbiology lab, were each assigned to conduct one of the methods. Developers of new methods provided the local users with a written Standard Operating Procedure (SOP) and basic training in the operation of the equipment prior to the exercise. These users also learned by observing, and assisting as appropriate, during the Phase I testing.

During both phases, all samples also were processed by five local laboratories (Table 1) using methods they employ in their routine processing. For enterococci, this included both Enterolert® chromogenic substrate (CS) (IDEXX, Westbrook, ME) and membrane filtration (MF) methods.

Data analysis

The primary means of data analysis was to compare the results from the new methods to that from traditional methods employed by the reference laboratories. This was done in three ways. First, we assessed the number of individual samples from each new method that differed by half a log unit from the reference laboratory median. Half a log unit was selected because previous laboratory intercalibration

Table 1. Local laboratories that analyzed the test samples using presently approved methods.

Laboratory	MF Method	CS Method
Los Angeles County Sanitation Districts	EPA 1600	Enterolert Quantitray® 2000
City of Los Angeles	Std. Methods 9230 C.	Enterolert Quantitray® 2000
Orange County Sanitation District	Did not perform MF	Enterolert Quantitray® 2000
Orange County Public Health Laboratory	EPA 1600	Enterolert Quantitray® 2000
City of San Diego	EPA 1600	Enterolert Quantitray® 2000

studies (Noble *et al.* 2004; Griffith *et al.* 2004) have demonstrated that this is the typical range of variability observed for traditional methodologies, both within and among laboratories. For this analysis, blank samples were counted as outside of range when values exceeded 50 cells/100 ml.

Second, results were individually evaluated for false positives and false negatives relative to the State of California standard of 104 cells/100 ml, as the State requires posting warning signs for any sample that occurs above this level. The decision of whether a sign should have been posted that day, against which the new methods were being evaluated, was based on the median concentration for that sample as measured by the reference laboratories.

The final analysis assessed precision of the measurements, which we could do because each sample was processed three times as blind replicates. Precision was quantified as the average coefficient of variation (CV) and was compared between the new methods and the reference laboratories. In all cases, the CV was calculated using the standard deviation and mean across all sample types for each method or reference laboratory. For purposes of data analysis, qualifiers (< or >) were ignored and only the numerical value was used.

Results

Q-PCR

Of the methods evaluated, Q-PCR was the most comparable to the reference labs, exhibiting the lowest rate of false positives (Figure 1) and the fewest number of samples more than 1/2 log unit from the reference laboratory median (Table 2). This method performed best on samples that contained moderate to high levels of enterococci in a simple seawater matrix without added interferences. However, Q-PCR tended to overestimate levels of enterococci, producing a higher mean value than traditional methods for two-thirds of the samples (Table 3).

Interferences from humic acids were not evident with Q-PCR, perhaps because the polycarbonate filters used do not retain them. Interferences from suspended solids were observed. With exception of Yorktown drain, results from the natural water samples showed good agreement between Q-PCR and reference method analyses (Tables 2 and 3).

Q-PCR exhibited relatively poor precision between replicate samples (Table 4), particularly when levels of enterococci were low. The greatest variability was observed for blanks (including filtered samples) and seawater samples containing interferences. Variability was lowest for samples containing a simple seawater matrix spiked with sewage.

Comparability to existing methods was about the same when Q-PCR was performed by local users in Phase II, as it was in Phase I. Q-PCR overestimated slightly relative to existing methods, but not by enough to cause false positives and within the measurement error typically encountered within existing methods (Table 5). Precision during Phase II was even better than when conducted by the Phase I method experts (Table 4).

Local practitioners who performed the Q-PCR method during Phase II found it to be relatively simple to complete. Sample analysis from start to finish took about 3 h for the 9 samples. Despite the speed with which the analysis was completed, several hours of preparatory work were required to perform tasks such as labeling tubes, preparing reagents, and setting up equipment. Participants felt that the amount and difficulty level of the preparatory tasks were similar to that required to perform standard membrane filtration analyses.

Dual Wavelength Fluorimetry

Comparability of DFW to the reference methods was sample dependent. The method exhibited a high

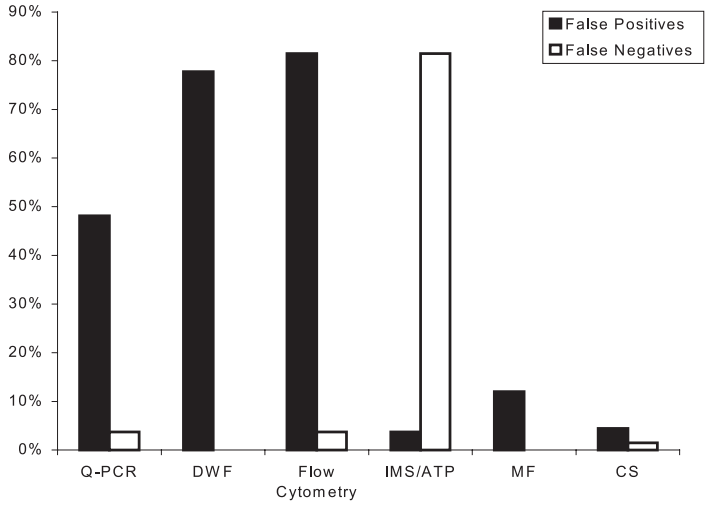

Figure 1. Percentage of false positives and false negatives with respect to California's warning threshold of 104 cfu/ 100 mL based on comparison to the reference laboratory grand median.

Table 2. Percentage of samples differing by more than fi log unit from the grand median produced by reference laboratories during Phase I.

Method	Seawater with Sewage n=9	Seawater with Urban Runoff n=9	Seawater with Interferences n=12	Natural Samples n=12	Total n=42
IMS/ATP	100	89	92	83	90
Flow Cytometry	67	100	25	42	62
DWF	11	100	25	50	45
Q-PCR	0	33	42	25	33

number of false positives (Figure 1), particularly for samples containing urban runoff either as an inoculum into seawater or as a natural sample. It was most comparable on samples consisting of a natural seawater matrix with moderate to high levels of bacterial contamination (Tables 2 and 3).

DFW displayed the best precision among the new methods tested. Variability of results for replicate samples during both Phase I and Phase II was comparable to that of the reference laboratories (Table 4).

Potential end-users that performed DFW during Phase II liked it because of its similarity to the chromogenic substrate methods they currently employ in their laboratories. However, they found that performing the method was labor intensive and could be somewhat tedious. The main reason was that the operator is required to transfer each sample back and forth between the water bath incubator and the measurement device multiple times per hour during the analyses, and so must remain close to the instrument. Users also observed that in its current configuration, the instrument required considerable set-up time prior to starting sample analysis. Despite this, the novice operators had a positive outlook on the future of the DWF method. They found the method to be straightforward to perform and felt that its operational shortcomings could be easily overcome by automating certain steps of the analysis protocol.

Flow Cytometry

The flow cytometry method consistently overestimated *Enterococcus* levels relative to present methods (Table 3), leading to a false positive rate relative to state standards of more than 50% (Figure 1). The problem appeared to be associated with high background counts, as values higher than state standards were measured in 11 of the 12 blank samples. This overestimation problem was also encountered in Phase II of the testing (Table 5).

Variability among replicate samples during Phase I was typically low and the method displayed superior precision except when presented with sterile phosphate buffered saline or urban runoff (data not shown). Variability during Phase II was comparable to that of the reference laboratories.

Potential end-users who performed this method during Phase II found it to be labor intensive. Of particular concern was the amount of time it took to complete all of the necessary steps in the protocol before the sample could be loaded into the instrument. As with DFW, there was a consensus among users that the method could be improved by automating many of the more labor-intensive steps.

IMS/ATP

IMS/ATP consistently underestimated levels of *Enterococcus* in the test samples (Tables 2 and 3), which resulted in a high rate of false negatives (Figure 1). This held true regardless of the different sample matrices, the presence of interferences, or whether the sample was natural or laboratory created. Variability among replicates in Phase I was high, but was similar to the reference laboratories during Phase II (Table 4).

Potential end-users liked the simplicity of the concepts behind this method and the fact that, unlike the other methods tested, it was designed for field use. However, they also found the method to be labor intensive and felt that some streamlining of the protocol would be necessary to adapt the method for processing multiple samples in the laboratory.

Discussion

Q-PCR was the most accurate of the methods, but it generally overestimated *Enterococcus* concentrations relative to the culture-based methods. This is consistent with previous Q-PCR applications in environmental water samples, where culture-based