Evaluation of the methods for enumerating coliform bacteria from water samples using precise reference standards

T. Wohlsen1,3, J. Bates1, G. Vesey2, W.A. Robinson3 and M. Katouli3

1 Public Health Microbiology, Queensland Health Scientific Services, Coopers Plains, Brisbane, Qld, Australia
2 BTF Pty Ltd, North Ryde, NSW, Australia
3 Faculty of Science, Health and Education, University of the Sunshine Coast, Maroochydore, Qld, Australia

Introduction

Coliform bacteria have been used for many years to determine the quality and safety of water for human consumption. *Escherichia coli* and other groups of coliforms may be present where there has been faecal contamination originating from warm-blooded animals (Chao et al. 2003). The presence of these bacteria in drinking water may indicate contamination resulting from a failure in the disinfection process (Tortorello 2003). However, the absence of these bacteria in water does not necessarily guarantee the absence of pathogens (Krewski et al. 2004). For example, Thurman et al. (1998) analysed several reservoirs and creeks in Australia and demonstrated no relationship between indicator bacteria and the detection of *Cryptosporidium* and/or *Giardia*.

Methods routinely used in water quality testing laboratories include; membrane filtration (MF) (Eckner 1998); most probable number (MPN) techniques (Grasso et al. 2000) and defined substrate technology (DST) systems (McFeters et al. 1993). The most commonly used methods in Australian laboratories include MF using membrane faecal coliform (mFC) agar and membrane lauryl sulphate (MSL) agar, standard plate count (SPC) including pour and spread plates, Colilert™ and Colisure™ that utilise nutrient indicators to simultaneously detect both total coliforms and *E. coli* (DST), the MPN technique and Petrifilm, a disposable plate impregnated with medium. Each procedure has an ideal target range for enumerating bacteria. For MF the range is 20–80 CFU, SPC 25–250 CFU, DST 0–200 CFU, MPN 2–1600 CFU and Petrifilm 15–150 CFU. Higher counts are obtainable using serial dilutions of the original sample.

In order to evaluate the efficacy of these methods, there is a need for a standard reference material with a well-defined number of bacteria. Freeze-dried bacterial cultures...
for proficiency testing have been utilised for many years (Peterz and Norberg 1983). However, the enumeration of these bacteria was traditionally performed by counting colonies recovered on agar plates. Recently a freeze-dried pellet that contains a precise number of viable bacteria has been developed (Morgan et al. 2004). The BioBall\textsuperscript{TM} (BTF Pty Ltd, NSW, Australia) is a white sphere, approximately 3 mm in diameter containing viable organisms produced by a cell sorting flow cytometer, programmed to count 30 organisms and deposit them into liquid nitrogen, which is then freeze-dried (Morgan et al. 2004). Bacterial cultures (30 organisms per ball) are supplied with a Quality Assurance Certificate of Analysis stating the mean and standard deviation (SD) of the batch.

In this study we used two BioBall\textsuperscript{TM} bacterial cultures to evaluate the precision of eight different methods for the enumeration of coliform bacteria in water samples using standard methods (Standards Australia 1995b; Clesceri et al. 1998) to ensure reproducibility.

**Materials and methods**

**Organisms and inoculation procedure**

BioBall\textsuperscript{TM} vials containing freeze-dried E. coli (ATCC = 11 775) and Enterobacter aerogenes (ATCC = 13 048) cultures were obtained (BTF Pty Ltd). A minimum of 10\% of each batch had been quality checked by the manufacturer using randomly selected vials. The quoted acceptance criteria for each batch was a mean of 28–33 colony-forming units (CFU) and a SD of < 3. BioBall\textsuperscript{TM} cultures, obtained from known reference strains, are four passages from the original strain and are accredited as a certified reference material according to the International Standard Organisation (ISO) Guide 34 and tested in an accredited laboratory to ISO/IEC 17 025 requirements. All methods were tested 10 times using each BioBall\textsuperscript{TM} culture for E. coli and Ent. aerogenes and the mean number ± SD of all readings was calculated. Three different batches were used.

**Membrane filtration**

Ten E. coli and 10 Ent. aerogenes BioBall\textsuperscript{TM} were each aseptically dropped to 100 ml autoclaved tap water and filtered onto a 0.45 μm mixed cellulose membrane (Millipore Australia Pty Ltd, NSW, Australia). Escherichia coli filters were placed onto mFC agar (Oxoid, Adelaide, Australia) and incubated at 30°C ± 0.5°C for 4 h, then at 44–44.5°C for 14 h according to the Australian Standard method (AS4276.7; Standards Australia 1995e). Escherichia coli was identified by the presence of typical blue colonies.

Enterobacter aerogenes filters were placed onto mLS agar (Oxoid) and incubated at 30 ± 1°C for 4 h, then at 37 ± 1°C for 14 h according to the Australian Standard method (AS4276.5; Standards Australia 1995b). Enterobacter aerogenes was identified by the presence of typical yellow/orange colonies.

**Most probable number**

The 15-tube multiple dilution technique for estimating the MPN of coliforms and E. coli was used. Ten E. coli and 10 Ent. aerogenes BioBall\textsuperscript{TM} were each aseptically dropped into separate 100 ml autoclaved tap water samples. Samples were inoculated into a set of five tubes containing improved formate-lactose-glutamate liquid medium (IFLG, Oxoid) at each required dilution (i.e. 5 tubes × 10 ml of sample, 5 × 1 ml of sample and 5 × 0.1 ml of sample). Tubes were incubated at 37°C for 48 h according to the Australian Standard method (AS4276.6 for E. coli and AS4276.4 for Ent. aerogenes; Standards Australia 1995a,d). Tubes were inspected for acid and gas production at 24 and 48 h and the combination of positive tubes were used to calculate the MPN of bacteria per 100 ml using the chart provided.

**Standard plate count**

**Pour plate**

Ten E. coli and 10 Ent. aerogenes BioBall\textsuperscript{TM} were each aseptically dropped into separate sterile Petri dishes and 1 ml of autoclaved tap water was added to dissolve the balls. Approximately 20 ml of molten (45°C) plate count agar (PCA, Acumedia, Melbourne, Australia) was poured into the plates (PCA-P), evenly distributed and incubated at 37°C for 48 h (Australian Standard method, AS4276.3.1; Standards Australia 1995c). Colonies were counted using an illuminated magnifying colony counter. To evaluate any difference between the use of water and saline, this procedure was repeated using sterile saline.

**Spread plate**

Spread plates using tryptone soya (TSA) agar, (Oxoid) and PCA were predried to remove 3 ml of moisture by incubating them overnight at 37°C. Ten E. coli and 10 Ent. aerogenes BioBall\textsuperscript{TM} were each aseptically dropped onto the surface of separate TSA (TSA-S) and PCA (PCA-S) plates and 1 ml of autoclaved tap water was added. Dissolved balls were spread evenly across the surface of the agar with an L-shaped spreader and incubated at 37°C for 48 h (Clesceri et al. 1998). Colonies were counted using an illuminated magnifying colony counter. This procedure was repeated using sterile saline.
Defined substrate technology

Colilert
Ten E. coli and 10 Ent. aerogenes BioBall™ were each aseptically dropped into separate 100 ml autoclaved tap water samples. Colilert 18 reagent (IDEXX Laboratories, NSW, Australia) was added and mixed well to dissolve. Samples were poured into a 51-well quanti-tray (IDEXX) providing counts from one to 200 per 100 ml of sample, sealed and incubated at 37°C for 24 h. Colilert simultaneously detects total coliforms by counting the number of yellow wells; and E. coli by counting yellow wells that also fluoresce at 365 nm. The MPN was calculated using a semi-automated quantification model provided by IDEXX.

Colisure
Ten E. coli and 10 Ent. aerogenes BioBall™ were each aseptically dropped into separate 100 ml autoclaved tap water samples. Colisure reagent (IDEXX Laboratories, NSW, Australia) was added and mixed to dissolve. Samples were poured into the 51-well quanti-tray, sealed and incubated at 37°C for 24 h. Colisure simultaneously detects total coliforms by counting the number of red/magenta wells; and E. coli by counting red/magenta wells that also fluoresce at 365 nm. Questionable wells (pink/orange) were incubated for up to 48 h. The MPN was calculated as mentioned previously.

Petrifilm
Ten E. coli and 10 Ent. aerogenes BioBall™ were each aseptically dropped onto the 3 M™ Petrifilm™ and 1 ml of sterile water was added to dissolve, spread evenly over the surface of the agar by pressing down with the instrument provided (Biolab Australia, Clayton, VIC, Australia) and incubated at 37°C for 48 h. Cultures was enumerated by counting the number of blue colonies (E. coli) and red colonies producing gas (Ent. aerogenes). This procedure was repeated using sterile saline.

Statistical analysis
For each method the mean number of recovered bacteria was compared with the values supplied for BioBall™ in that specified batch using a single sample t-test. The sample SD was tested against the specified SD reported by the BioBall™ for each batch using the chi-square distribution (Zar 1984). Because some batches had different specifications (i.e. mean and SD), the number of recovered bacteria from each test was converted to a standardised value by subtracting the recovered number of bacteria from the specified mean of the batch used. A three-way factorial ANOVA was then used to determine whether there was a difference in the mean standardised values recovered between bacteria, between methods and between using saline and water (PCA-P, TSA-S, PCA–S and Petrifilm). Interactions were followed up using separate ANOVA for each bacterium and graphs were used to interpret significant interactions.

Results
BioBall™ batches used in these experiments had specified means ranging from 29·3 to 30·9 organisms for E. coli and 30·4 to 31·8 organisms for Ent. aerogenes. SD ranged from 2·4 to 2·9 for E. coli and 2·5 to 2·9 for Ent. aerogenes. In our study, the mean recovery of both E. coli and Ent. aerogenes varied with the different methods employed. For E. coli the mean recovery ranged between 15·9 ± 7·13 for MF and 31·6 ± 5·93 for Colilert (Table 1). For both bacterial strains tested, the best and most consistent results were obtained with Petrifilm (28±6 ± 2·01 for E. coli and 29·8 ± 1·81 for Ent. aerogenes) followed by the TSA-S (28·3 ± 1·64 for E. coli, 28·5 ± 1·58 for Ent. aerogenes). Petrifilm and PCA-P methods always returned means and SD within specification but for E. coli, the MF, PCA-S and Colisure methods recovered significantly lower mean counts than specified. MF, Colilert, Colisure and MPN methods had significantly higher variability between replicates than expected (Table 1, Fig. 1). For Ent. aerogenes, MF and TSA-S methods recovered significantly lower mean counts than specified, while Colilert, Colisure and MPN methods again had significantly higher variability between replicates (Table 1, Fig. 1). The three-way ANOVA indicated the effect of type of solution on the recovery values by each method depended on the bacteria used (i.e. Fbacteria/method/solution = 3·62, d.f. = 3,144, P < 0·05). ANOVA for E. coli alone found the effect of method depended on substrate (Fmethod/solution = 3·08, d.f. = 3,73, P < 0·05) with saline yielding a lower recovery for PCA-P, TCA-S and Petrifilm (Fig. 2). For E. aerogenes, the interaction between method and substrate was also significant (Fmethod/solution = 4·82, df = 3,72, P < 0·005) with saline yielding lower recovery for all methods except TSA-S (Fig 2). Using water all four methods yielded a mean recovery within two of the specified values (mean = 1·28) for all tests with the exception of PCA-S, which yielded an average count four below the specified value for E. coli. On average, methods using saline yielded counts of 2·7 below specified values (Fig. 2).
**Table 1** Recovery of eight methods using BioBall™ for each bacteria (n = 10). Only significant P-values are given

<table>
<thead>
<tr>
<th>BioBall™ organism</th>
<th>SPC</th>
<th>MF</th>
<th>PCA</th>
<th>TSA</th>
<th>PCA</th>
<th>DST</th>
<th>Colilert</th>
<th>Colisure</th>
<th>MPN</th>
<th>Petrifilm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Escherichia coli</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted mean per batch</td>
<td>29.3</td>
<td>29.3</td>
<td>29.3</td>
<td>29.3</td>
<td>29.3</td>
<td>29.3</td>
<td>30.9</td>
<td>30.9</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Probability recovered mean</td>
<td>15.9</td>
<td>27.8</td>
<td>28.3</td>
<td>25.7</td>
<td>31.6</td>
<td>16.6</td>
<td>29.3</td>
<td>28.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value (t-test)</td>
<td>0.0002</td>
<td>&lt;0.0001</td>
<td></td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability recovered SD</td>
<td>7.13</td>
<td>2.44</td>
<td>1.64</td>
<td>1.49</td>
<td>5.93</td>
<td>6.15</td>
<td>9.73</td>
<td>2.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value (χ² test)</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation %</td>
<td>44.8</td>
<td>8.8</td>
<td>5.8</td>
<td>5.8</td>
<td>18.8</td>
<td>37.0</td>
<td>33.3</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Enterobacter aerogenes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted mean per batch</td>
<td>31.8</td>
<td>30.4</td>
<td>30.4</td>
<td>30.4</td>
<td>31.8</td>
<td>31.8</td>
<td>31.8</td>
<td>30.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability recovered mean</td>
<td>27.3</td>
<td>30.3</td>
<td>28.5</td>
<td>29.7</td>
<td>33.6</td>
<td>31.7</td>
<td>35.6</td>
<td>29.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value (t-test)</td>
<td>0.0240</td>
<td>0.0042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability recovered SD</td>
<td>5.25</td>
<td>2.36</td>
<td>1.58</td>
<td>1.95</td>
<td>5.38</td>
<td>4.45</td>
<td>10.24</td>
<td>1.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-value (χ² test)</td>
<td>0.0074</td>
<td></td>
<td></td>
<td></td>
<td>0.0003</td>
<td>0.0119</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation %</td>
<td>19.2</td>
<td>7.8</td>
<td>5.5</td>
<td>6.6</td>
<td>16.0</td>
<td>14.0</td>
<td>28.8</td>
<td>6.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MF, membrane filtration; SPC, standard plate count; DST, defined substrate technology; MPN, most probable number; PCA, plate count agar; TSA, tryptone soya agar.

Italicized value indicate SD was less than specification, bold values indicate SD greater than specification.

**Discussion**

Generally, similar studies lack a standard reference inoculum to compare methods, with bacterial inoculum enumerated by serial dilution normally being inconsistent. We used a precise reference standard to maintain the inoculum size and this provided us with a better assessment of these methods for enumerating bacteria. Both selective (MF, DST, MPN and Petrifilm) and nonselective (SPC) techniques were chosen to assess the performance of the BioBall™ bacteria.

Grasso et al. (2000) reported that MPN counts were consistently higher than Colilert counts. Similar results have been found comparing MF and Colilert (Fricker et al. 1997). In our study, MPN overestimated the concentration for Ent. aerogenes compared with other methods. However, using the coefficient of variation we found that not only MPN, but also Colisure and MF methods showed high variability. Chao et al. (2003) showed that mFC agar yielded lower counts to enumerate E. coli than Colilert and that 18% of samples failed to produce the ‘typical’ blue colonies in waters from sub-tropical areas. We also obtained low average counts using mFC and possible reasons for this could be the high selectivity of the mFC medium and the higher incubation temperature (44°C) compared with other methods (37°C). These bacteria have also been subjected to liquid nitrogen and freeze drying before testing; therefore the organisms may be stressed resulting in the lower recovery rate of the bacteria at this higher incubation temperature, despite the 30°C resuscitation step. We chose agars and incubation temperatures to comply with reference methods.

McFeters et al. (1993) demonstrated that Colilert and Colisure produced comparable results to MF and MPN procedures using chlorine-stressed organisms and that Colisure produced the highest counts between 28 and 48 h incubation. We required 48 h incubation and demonstrated that Colisure and mFC yielded the lowest recovery rate for E. coli compared with other methods, where other studies have shown good reproducibility between the DST and MPN methods (Edberg et al. 1991). Both Colilert and Colisure use β-glucuronidase to produce fluorescence from 4-methylumbelliferyl-β-d-glucuronide (MUG) to identify E. coli. To identify coliforms, Colilert combines o-nitrophenyl-β-d-galactopyranoside (ONPG) as a substrate for β-galactosidase to produce a
yellow colour, while Colisure uses β-galactosidase to metabolise chlorophenyl red (CPRG) to produce magenta. As identification of E. coli requires fluorescence and a colour change, it is possible this variation in formulation affects the recovery of E. coli with Colisure as seen in our study. The high range of recovery rate in 10 replicate samples for both bacterial strains tested did not indicate a high reproducibility.

Standard plate count, in particular the pour plate method, has been used for estimating bacterial loads of water samples for many years. It has been postulated that the addition of molten agar (45°C) to water samples can cause stress to the microorganisms, resulting in a decreased recovery (Reasoner 2004). This however, was not demonstrated by our results where the recovery of both organisms was slightly higher using PCA-P than PCA-S (Table 1).

We compared recovery rate and SD of the methods to show that while some had a SD within the accepted level of BioBall™, some had a mean recovery rate within the accepted level with high a SD. Increased variability can result in less power in some statistical analysis and cloud
over significant results, similarly a high SD indicates a low reliability and reproducibility of the method.

With a large number of hypothesis tests performed on the same batch of BioBall™ one could expect some significant effects because of type I error, i.e. false positives simply by chance alone with a large number of tests. However the majority of significant effects in this study were at the level of 0.005 or less, suggesting that type I error was not a cause of the high number of significant results.

For each method we also calculated a coefficient of variation for assessment of relative variability of the methods. MF and Colisure methods were the most variable methods for both bacteria and had significantly higher SD than specified by BioBall™. Their use may therefore result in low power in statistical tests. Colisure returned high average counts (i.e. 41 and 42.7) but neither of which were detected as significant by the t-test.

Using the sample size formula of Zar (1984) we calculated that 251 replicates of E. coli and 278 replicates of Ent. aerogenes would be required to obtain a mean with ± 2 using the MPN procedure. While 102 replicates of E. coli and 54 replicates of Ent. aerogenes would be required to obtain a mean within ±2 using the Colisure procedure. For the MF procedure, 136 replicates of E. coli (mFC) and 75 replicates of Ent. aerogenes (mLS) would be required to obtain a mean within ±2. The Petrifilm (for E. coli n = 11; Ent. aerogenes n = 12), PCA-P (18;17), TSA-S (9;9) and PCA-S (8;12) methods would be considerably more efficient.

Escherichia coli BioBall™ generally produced lower recovery using all methods than Ent. aerogenes BioBall™ indicating that the type of coliform bacteria employed as well as the method might influence their percentage recovery in water samples. We also found that BioBall™ re-hydrated in water produced recoveries closer to the specified values for both E. coli and Ent. aerogenes than those re-hydrated in saline. Only in two instances did saline produce higher recoveries than water, PCA-S (E. coli) and TSA-S (Ent. aerogenes). This indicates that osmotic shock because of re-hydration in water instead of saline does not affect the recovery of these BioBall™ bacteria.

In conclusion, for the methods enumerating 1-ml samples (Petrifilm and SPC), Petrifilm and PCA-P produced the most accurate and consistent results. From the methods specifically targeted to isolate coliform bacteria from 100-ml water samples (MF, DST and MPN), Colilert produced the most consistent results with the lowest coefficient of variation. Our data also indicated that methods routinely used by laboratories such as MF, Colisure and MPN for water quality testing may be producing results with greater variability than first thought. However, the application of these methods for recovery of other species of coliform bacteria should be evaluated before these results can be generalized. We also conclude that the BioBall™ is a very suitable quality control tool for evaluating the efficiency of methods for bacterial enumeration in water samples.

Acknowledgements

We thank Bruce Gray, Marie Napoli, Leonie Cover and Trudy Graham for their technical assistance and helpful comments. We also thank the staff of the PHM Laboratory and the QHPSS Library for their invaluable assistance.

References


