SUMMARY

This document describes a template protocol for laboratory assessment of field test kits which could in the future be purchased by WHO and UNICEF to support water quality testing activities of national partners. It has been developed with reference to other evaluations and assessments, especially the ISO standards 17994 and 16140-2, the US EPA's Alternative Test Procedure and a published evaluation of different total coliform and *E. coli* tests (Olstadt, 2007). See the Bibliography for a summary of these resources.

The template protocol should be reviewed with the laboratory responsible for conducting the tests and adapted as necessary to the particular characteristics of the assay under evaluation. For each assay under evaluation, a technology-specific testing protocol will be developed based on the template protocol. The template protocol describes two phases of assessment: a first phase with a limited number of tests (approximately 20-40) to determine if the trial method produces results comparable to a reference method for *E. coli* under standard conditions; and a second phase with approximately 130 tests, to determine if the trial method is robust with respect to false negatives and positives, in different water matrices and (if appropriate) at different incubation temperatures. The trial method is compared against a reference method (20-40 tests for Phase 1, and 21 tests for Phase 2) using a range of challenge waters meeting specific quality guidelines. If the same challenge waters are used, one set of reference results may be used for comparison against multiple trial methods tested at the same time.

A trial method must meet minimum thresholds of concordance with the reference method in order to move on to Phase 2 testing. In Phase 2 testing trial methods will not be given "pass/fail" scores, but rather will be assessed against different criteria, and the degree of compliance for each criterion will be noted in a final test report.

BACKGROUND

The primary concern regarding drinking water quality is that faecal contamination of drinking water could lead to disease. A large number of pathogens can cause water-borne disease, including viruses, bacteria, and protozoa. The majority of pathogens causing water-borne diseases are faecal in origin, but it is not practical to test drinking water for all potential pathogens. Instead, measurement of faecal indicators is preferred. While all faecal indicators have limitations, there is widespread agreement that *Escherichia coli* (*E. coli*) is the best currently available indicator of faecal contamination in drinking water, and that thermotolerant coliforms are an acceptable alternative[1]. Many of the portable water quality testing equipment sets currently available measure *E. coli* or thermotolerant coliforms in drinking water, though some innovative technologies make use of alternative measures.

There are a number of standard methods for enumeration of coliform bacteria including *E. coli*[2,3,4,5], but these can be difficult to apply outside of laboratories. Especially in remote areas there is a need for portable field kits which can permit a trained technician to measure water quality in the field, including in areas lacking electricity. UNICEF and WHO currently procure significant numbers of water quality testing kits, as do other UN

[1] Guidelines for Drinking Water Quality, https://www.who.int/water_sanitation_health/publications/drinking-water-quality-guidelines-4-including-1st-addendum/en/
[2] ISO 4832, 7251, 9308, and 16649.
[3] Standard Methods for the Examination of Water and Wastewater, methods 9221, 9222, and 9223.
[4] US EPA Methods 1603 and 1604
[5] AOAC official methods 991.14, 998.08, and 110402.

agencies and NGOs, particularly in the emergency response sector. However, there is currently a lack of protocol to assess such the quality and accuracy of these portable field testing kits.

The objective of the proposed laboratory assessment is to evaluate portable field testing kits under a variety of conditions, based on the claims of the manufacturer, and compare results against standard methods for measurement of *E. coli* in drinking water. Each assay under evaluation will be assessed according to a standard set of conditions, as well as to relevant specific conditions claimed by the manufacturer (e.g. incubation temperatures and times). This round of assessments will focus on culture-based assays which aim to detect faecal indicator bacteria; subsequent rounds may include more novel proxies of faecal contamination that do not rely on incubation, e.g. nucleic-acid assays or real-time monitoring sensors.

Assessments will be made in two phases: a first phase to assess if the assay under evaluation accurately measures low, medium, and high concentrations of *E. coli*; and a second phase to assess the performance of the assay with respect to false negatives, false positives, competing organisms, different matrices, and different incubation temperatures.

In both phases, a number of principles will be followed:

1. Assays under evaluation will be used to test a series of challenge waters which will be prepared with varying levels of *E. coli*, prepared from a fixed laboratory strain of *E. coli* in a matrix of phosphate buffered saline. In Phase 2 wild strains of *E. coli* will also be used.
2. Samples will be incubated at the temperature and for the recommended amount of time claimed by the manufacturer. If the manufacturer claims that fixed-temperature incubation is not required, two sets of samples will be analysed and incubated at 20°C and 35 °C. Results will be recorded at any time specified by the manufacturer, as well as at 24 hours and 48 hours (±4 hours).
3. A specified number of replicate tests and blank tests will be conducted for each condition being tested.
4. If assays involve hardware which is re-used (e.g. filtration apparatus) three separate sets of equipment will be used in the assessment in order to assess the variability between instruments. If possible, the three sets of equipment should be drawn from different production lot numbers.

Technology-specific protocols will be developed for assays under evaluation based on these principles and the assessment procedures described in following sections. To the extent possible the same test waters will be used for the different assays, and the same reference method will be used (SM 9223B, IDEXX Colilert-18 Quanti-Tray 2000 system).

In addition to the standard conditions described in this document, additional conditions could be added if there are particular claims by the manufacturer.

PHASE 1 ASSESSMENTS

The first phase aims to determine if the assay under evaluation produces results comparable to the reference method over a range of *E. coli* concentrations, under highly controlled conditions.

A stock solution of *E. coli* from a known lab strain (ideally, ATCC 25922, but to be confirmed with the laboratory) will be prepared with a concentration of approximately 1000 viable and culturable *E. coli* cells per 100 mL[6] (acceptable range: 300 – 3000 cells/100 mL, to be measured using the IDEXX Quantitray method) in a background of sterile phosphate buffered saline (pH 7.4 ± 0.2). This stock solution will then be serially diluted using two-fold dilution with sterile phosphate buffered saline, as per the table below. The resulting stock

[6] Cell density can be measured either by counting Colony Forming Units (CFU), or with a statistical Most Probable Number (MPN) method. The reference method (IDEXX Quantitray) gives results in MPN/100 mL.

solutions will span a range of concentrations which would be expected to yield positive results ranging from zero to above most detection limits, with several stocks in critical ranges in between.

A sample of Stock 1 will be autoclaved and used as a blank, free from viable and culturable *E. coli*.

Table 1: Stock dilutions

Stock	Approximate *E. coli* concentration, cells/100 mL		
	Lower acceptable limit	Target concentration	Upper acceptable limit
S1	300	1000	3000
S2	150	500	1500
S3	75	250	750
S4	38	125	375
S5	19	64	188
S6	9	32	94
S7	5	16	47
S8	2	8	23
S9	1	4	12
S10	0.6	2	6
S11	0.3	1	3
S12	0.1	0.5	1.5

PRESENCE / ABSENCE TESTS

Some assays under evaluation will be presence/absence tests that do not attempt to quantify the concentration of faecal indicator bacteria. Presence/absence tests in most cases involve mixing a growth medium with a fixed volume of sample water, incubating for an appropriate time at a designated temperature, and checking for the development of colour or fluorescence.

The assays under evaluation will be used according to their instructions with each of six of the challenge stocks (S1, S3, S5, S7, S9, and S11) as well as the Blank (autoclaved S1) and incubated in 96-well deep well plates with at least 1 mL volumes (e.g. CLS3960 Sigma or similar, plates should be autoclavable for reuse). If the media require anoxic conditions an IDEXX Quantitray-2000 system can be used instead.

Samples will be incubated at the temperature recommended by the manufacturer and for the recommended amount of time. If the manufacturer claims that fixed-temperature incubation is not required, two sets of samples will be analyzed and incubated at 20 °C and 35 °C. Results will be recorded at any time specified by the manufacturer, as well as at 24 hours (±4 hours). For a secondary analysis, the same samples will be analyzed after 48 hours (±4 hours) to see if results vary significantly in case they are not recorded promptly. Results from the reference method will only be recorded at 24 hours.

For each condition, 3 replicate tests will be conducted for both the trial and reference method.

Table 2: Phase 1 Presence/Absence tests

Condition	Number of tests (trial method)	Number of tests (reference)
Varying concentrations of *E. coli*	6 x 3	6 x 3
Blanks	1 x 3	1 x 3
TOTAL	**21**	**21**

Each of the 96 wells will be recorded as having a positive or negative result, and the number of positive results will be used to generate a Most Probable Number (MPN) estimate of *E. coli* density. Results will be compared against the reference method, using the MPN approach. Since MPN analysis produces point estimates with

confidence intervals[7], the trial method will be considered to match the reference method when the 95% confidence intervals overlap.

Example: if the Colilert Quantitray 2000 test yields positive results in 40 large and 2 small wells, the Most Probable Number is 78 cells per 100 mL, with a 95% confidence interval of (56, 110). If the trial media yields 48 positive wells from the 96-well tray, the Most Probable Number is 69, with a 95% confidence interval of (52, 93). The confidence intervals overlap so the trial media is considered to agree with the reference method. But if the trial media yields only 30 positive wells the Most Probable Number is 37, with a 95% confidence interval of (26, 54). Since the confidence intervals do not overlap, the trial media would be considered to not match the reference method.

If the MPN estimates agree for at least five of the six stock solutions, the assay under assessment will proceed to Phase 2 testing. If any wells filled with the sterile blank are found to be positive with the trial media, the test will be redone. If any wells are found to be positive on this further test, while all the reference method wells with blanks are negative, the assay under assessment will not proceed to Phase 2 testing.

SEMI-QUANTITATIVE TESTS

Semi/quantitative tests report results in ranges of concentrations. This protocol assumes that 100 mL of sample is to be tested, and that results will be grouped into the following four risk classes:

A.	<1 CFU/100 mL
B.	1-10 CFU/100 mL
C.	11-100 CFU/100 mL
D.	> 100 CFU/100 mL

If other classes are proposed by the manufacturer the protocol will be adjusted accordingly.

Each of the 12 stocks from Table 1 will be tested, as well as the blank (autoclaved S1) using three separate trial kits. Each stock and the blank will also be measured using the reference method.

Samples will be incubated at the temperature recommended by the manufacturer and for the recommended amount of time. If the manufacturer claims that fixed-temperature incubation is not required, two sets of samples will be analyzed and incubated at 20°C and 35 °C. Results will be recorded at any time specified by the manufacturer, as well as at 24 hours (±4 hours). For a secondary analysis, the same samples will be analyzed after 48 hours (±4 hours) to see if results vary significantly in case they are not recorded promptly.

For each condition, 3 replicate tests will be conducted for both the trial and reference method.

Table 3: Phase 1 Semi-quantitative tests

Condition	Number of tests (trial method)	Number of tests (reference)
Varying concentrations of *E. coli*	12 x 3	12 x 3
Blanks	1 x 3	1 x 3
TOTAL	**39**	**39**

[7] E.g. Jarvis, B., Wilrich, C., and P.-T. Wilrich: Reconsideration of the derivation of Most Probable Numbers, their standard deviations, confidence bounds and rarity values. Journal of Applied Microbiology 109 (2010), 1660 – 1667. http://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/professoren/wilrich/MPN_ver6.xls

Results will be assessed by comparing the proportion of results that are in the same risk class, and the proportion that are within one risk class as measured with the reference method (IDEXX Quantitray 2000).

Example 1 : Perfect correlation, 100% (39) in same class.

		Reference Method results			
		A	B	C	D
Trial Method	A	9			
	B		9		
	C			9	
	D				12

Example 2 : Good correlation, 79% (31) in same class, 100% within one class

		Reference Method results			
		A	B	C	D
Trial Method	A	7	2		
	B	2	7	1	
	C		1	8	3
	D				9

Example 3 : Moderate correlation, 62% (24) in same class, 90% within one class

		Reference Method results			
		A	B	C	D
Trial Method	A	6	3	1	
	B	3	6	2	3
	C			6	3
	D				6

Example 4 : Poor correlation, 51% (20) in same class, 86% within one class

		Reference Method results			
		A	B	C	D
Trial Method	A	5	4	3	1
	B	4	5	1	1
	C			5	3
	D				5

An assay under assessment will be considered to match the reference method sufficiently when at least 60% of results (24 tests) are in the same class, and at least 90% (35) are within one class. In the previous examples, the first three demonstrate sufficient correlation to pass the Phase 1 assessment, though the third example is the poorest possible passing result. The fourth example would not have high enough agreement and would not proceed to Phase 2 testing.

QUANTITATIVE TESTS

Several kinds of assays under evaluation yield quantitative results. Some yield a point estimate with confidence intervals (e.g. MPN methods or colony counting methods). Other may yield a point estimate without confidence intervals (e.g. methods that measure nucleic acids, enzymes, or other compounds associated with faecal contamination).

Each of the 12 stocks from Table 1 will be tested, as well as the blank (autoclaved S1) using three separate trial kits. Each stock and the blank will also be measured once using the reference method.

Samples will be incubated at the temperature recommended by the manufacturer and for the recommended amount of time. If the manufacturer claims that fixed-temperature incubation is not required, two sets of samples will be analyzed and incubated at 20 °C and 35 °C. Results will be recorded at any time specified by the manufacturer, as well as at 24 hours (±4 hours). For a secondary analysis, the same samples will be analyzed after 48 hours (±4 hours) to see if results vary significantly in case they are not recorded promptly.

Table 4: Phase 1 Quantitative tests

Condition	Number of tests (trial method)	Number of tests (reference)
Varying concentrations of *E. coli*	12 x 3	12 x 3
Blanks	1 x 3	1 x 3
TOTAL	**39**	**39**

Results will be assessed by making linear regressions of the results from the assay under evaluation against the reference method results, following log transformation of both datasets. Within a given stock, triplicate samples from the assay under evaluation will be 'paired' with triplicate analyses made using the reference method during sample processing (before the incubation period).

Samples above the maximum detection limit will be fixed at the maximum detection limit, and samples below the minimum detection limit will be fixed at 50% of the detection limit. If the upper quantification limit (UQL) or the lower quantification limit (LQL) of the assays vary significantly from those of the reference method, a set of linear regressions will be made – first including all data points, and next including only data points that are within the quantification range for both assays. Two statistical tests will be made on the regression(s):

- The slope should not be significantly different from unity ($p<0.05$)
- The goodness of fit, as measured by Pearson's rank coefficient, should be at least 0.90.

An assay will proceed to Phase 2 assessment if both of the above conditions are fulfilled, and if the blanks do not show positive results.

PHASE 2 ASSESSMENTS

Assays that have passed Phase 1 assessments can proceed to the second phase of assessment, which will examine the performance of the test with respect to:

1. False positives due to non-target bacteria
2. False negatives due to competition from non-target bacteria
3. Incubation temperature
4. Water matrix and wild *E. coli*

Some of these tests may reveal limitations of an assay that would not necessarily lead to the assay being rejected: for instance if the assay under evaluation does not perform well with certain matrices it may still be suitable in other matrices.

The second phase of testing should ideally make use of the same *E. coli* stocks prepared for use in Phase 1, so if possible Phase 2 tests should be conducted within 24 hours of reading results from Phase 1.

The same protocols for Phase 2 testing will be followed for presence/absence, semi-quantitative and quantitative tests.

FALSE POSITIVES DUE TO NON-TARGET BACTERIA

Some tests may generate positive results in the absence of *E. coli*. This may be caused by growth of non-target organisms. Cultures of six non-target bacteria (*Aeromonas*, *Citrobacter*, *Enterobacter*, *Klebsiella*, *Pseudomonas aeruginosa* and *Serratia*) that could potentially cause false positive results will be produced to a target concentration of 100,000,000 viable and culturable cells/100 mL (acceptable range: 30,000,000 – 300,000,000 cells/100 mL), and tested using the trial assay, without addition of any *E. coli*. Any positive results will be considered as a false positive.

Single tests will be made (i.e. no triplicates) and the reference method will not be challenged with the non-target organisms.

FALSE NEGATIVES DUE TO COMPETITION

The same six cultures of non-target organisms will be mixed 1:1 with *E. coli* Stock 1, yielding an approximate 100,000:1 ratio of non-target organisms to *E. coli*. The resulting stock will be tested using the trial kit. Any negative results will be considered as an indication that the trial method may not detect *E. coli* in the presence of competing bacteria.

Single tests will be made (i.e. no triplicates) and the reference method will not be challenged with the non-target organisms.

TEMPERATURE

This condition will apply only to trial methods for which the manufacturer claims that fixed-temperature incubators are not required (e.g. incubation is recommended at ambient temperature).

The *E. coli* stock that yielded the count closest to 30 cells/100 mL (with the reference method) in Phase 1 testing will be used for testing at various temperatures. Six samples of this stock will be tested using the trial method, and incubated at the following temperatures: 20, 25, 30, 35, 40, and 45 °C. Results will be read and recorded at 24, 48, and 72 hours, and any additional times recommended by the manufacturer.

For each temperature, 3 replicate tests will be conducted for the trial method. A single reference method analysis will be made at the standard incubation temperature (35 °C) and 24 hours for comparison.

NATURAL WATERS

The water matrix as well as the strain of *E. coli* used may affect the performance of the trial method. To assess this possibility, five different natural waters will be selected, including at least two surface water and at least two groundwater sources. The general chemistry of these waters (pH, alkalinity, hardness, conductivity, turbidity) will be characterized, and at least one source should have turbidity >10 NTU, one source should have a pH of below 6.5, and one source should have a pH of greater than 8. At least one water should have low alkalinity (<50 mg/L as $CaCO_3$).

Ideally, the pH for all waters will be natural, but it is acceptable to add acid or base if a suitable natural sample can’t be found. Turbidity should not be increased artificially for sample N1.

Stock	Source: ground water (GW) or surface water (SW)	Turbidity	pH
N1	GW or SW	>10	Any
N2	GW or SW	<10	<6.5
N3	GW or SW	<10	>8.0
N4	GW or SW	Any	6.5-8.0
N5	GW or SW	Any	Any
TOTAL	At least 2 GW, at least 2 SW		

The natural waters will be autoclaved and then spiked with effluent from a wastewater treatment plant to reach a target concentration of 300 *E. coli* / 100 mL (acceptable range 100 – 1000). This may require some pre-testing to identify an appropriate dilution factor for the effluent. This stock solution will then be serially diluted using ten-fold dilution with the autoclaved natural waters three times. The resulting stock solutions will span a range of concentrations which would be expected to yield at least one stock in each of the risk classes A, B, C, and D. A sterile stock of the autoclaved natural water will be used as a fifth stock.

Stock	Approximate *E. coli* concentration, cells/100 mL		
	Lower acceptable limit	Target concentration	Upper acceptable limit
NxS1	100	300	1000
NxS2	10	30	100
NxS3	1	3	10
NxS4	0.1	0.3	1
Autoclaved	0	0	0

Each of the natural water stocks will be tested with the test method in triplicate, using three different sets of equipment. (5 waters * 5 stocks * 3 replicates using different equipment), and with the reference method (one test per stock). Comparisons will be made as in Phase 1.

REPORTING

For each trial method assessed, a final testing report will be produced, according to a structure agreed between the testing laboratory and WHO. The Phase 2 results will not be presented in terms of pass/fail, but rather the degree of compliance for each criterion assessed will be reported. An annex will include all raw data.

SUMMARY

Phase two testing will require 105 tests of the trial method (or 87, if the temperature assessments are not made) and 26 of the reference method.

Table 5: Number of tests in Phase 2

Condition	Number of tests (trial method)	Number of tests (reference)
False positives (non-target bacteria)	6	n.a.
False negatives (non-target bacteria)	6	n.a.
Temperature	6 x 3	1
Matrix, wild *E. coli*	5 x 5 x 3	5 x 5 x 1
TOTAL	**105**	**26**

LIMITATIONS

In order to keep the assessment fairly light, a number of factors will not be considered.

- This assessment will not include confirmations (for example using a 3rd method or API 20E testing).
- This assessment will not include the testing of a wide range of water sources (beyond the five natural waters) to determine performance of the tests for different types of water sources.
- This assessment will be laboratory-based and will not include assessments of strains from different contexts (e.g. tropical).
- This assessment will not include injured bacteria, for example chlorine injured bacteria.
- This assessment will focus on the technical performance of the trial kits rather than on their cost or user-friendliness. However, an informal assessment of ease of use will also be made by the analysis team members.

BIBLIOGRAPHY

This document draws upon several prior studies or programmes for assessment of microbiological testing procedures, the key features of which are summarized below.

ISO 17994:2014 (Water quality -- Requirements for the comparison of the relative recovery of microorganisms by two quantitative methods). https://www.iso.org/standard/56617.html

- Sets out a statistical basis for comparing two quantitative methods of measurements of microbial water quality, based on the paired *t*-test.
- The comparison considers two methods to be quantitatively "not different", or "different" based on the results of the *t*-test. If the variability is too high, the comparison may be considered "inconclusive". It's also possible that the methods are statistically different, but that the difference is too small to be of any practical concern: in this case the methods are called "indifferent".
- The standard compares two methods, neither of which is necessarily a reference method.
- A range of water types should be tested, with a preference given to natural samples. Spiking with pure cultures is considered a last resort.
- It is recommended to involve multiple laboratories in the comparison, but it's also possible that one laboratory can conduct a suitable comparison study provided they have access to a wide range of sample types.
- The number of samples required for a valid comparison can't be determined in advance. However, if the two methods are markedly different, a small number of samples could be sufficient. Accordingly it is recommended to proceed in stages.

ISO 16140-2:2016 (Microbiology of the food chain -- Method validation -- Part 2: Protocol for the validation of alternative (proprietary) methods against a reference method). https://www.iso.org/standard/54870.html

- Provides a protocol for comparing a novel method against a reference method, for quantifying microbiological contamination of food products
- Comparison can be made in paired or unpaired studies.
- Specifies minimum numbers of samples (at least 60 composed of at least three types of food)
- Calls for an interlaboratory study involving at least 10 laboratories, and a minimum of 480 results per laboratory.

US EPA Microbiological Alternate Test Procedure (ATP) Protocol for Drinking Water, Ambient Water, Wastewater, and Sewage Sludge Monitoring Methods. https://www.epa.gov/sites/production/files/2015-09/documents/micro_atp_protocol_sept-2010.pdf

- Reference methods: LTB-EC MUG (MPN), nutrient agar with *E. coli* (MF)
- Comparability study: Drinking water (oxidant-free) x 1, Wastewater x 10 (20 replicate each),
- Specificity study: 100 x 2 (FPs and FNs), 200 x 2 for quantitative methods
- Strong reliance on confirmations
- Control strains: None, wastewater used throughout. "If samples are spiked, environmental isolates should be used, as pure strains may exhibit different recovery and precision characteristics than natural flora. NELAC (Reference 10.9) and ATCC (Reference 10.6) recommend that bacterial cultures be transferred monthly and passed no more than five times before returning to the original culture."
- Analysis: Mean recovery for each matrix, precision of estimates (relative standard deviation), sensitivity and specificity (relative to independent standard not EPA approved reference method) with a chi-2 test for difference and Breslow-day for differences across matrices. For quantitative tests a flow chart explains which tests to use depending on whether the data are normal or log-normal.
- Note: There is a whole range of QC measures that laboratories conducting studies must follow

Olstadt et al. 2007. A comparison of ten USEPA approved total coliform/*E. coli* tests. https://pubmed.ncbi.nlm.nih.gov/17674575/

- A comparative study designed to assess performance of a number of assays in measurement of total coliforms.
- Reference methods: None
- Samples: 3 sampling sites (2 x), Low (1-10) and high (50-100) spike levels. A total of 1100 samples across 10 tests
- Control strains: 5 total coliform strains (*E. coli, Klebsiella, Enterobacter, Citrobacter* and *Serratia*) + 2 aeromonas strains, isolated from drinking water samples, suppression up to 10^6 aeromonas
- Analysis: P/A % failure rate to detect TC/*E. coli*, quantitative (% recovery vs HPC counts from triplicate experiments)

Genter et al. 2019. Evaluation of the novel substrate RUG™ for the detection of *Escherichia coli* in water from temperate (Zurich, Switzerland) and tropical (Bushenyi, Uganda) field sites. https://pubs.rsc.org/en/content/articlehtml/2019/ew/c9ew00138g

- Paper resulting from MSc work of Franziska Genter, supervised by Dr Tim Julian and Dr Sara Marks (Eawag) with technical support from the JMP team
- Comparison of resorufin media to standard reference methods
- Reference methods: Colilert QT, mTEC (Macconkey and API 20E used for confirmations)
- Control strains: The positive control strain was *E. coli* (ATCC 25922) and the negative control strains were *Enterococcus faecalis* (ATCC 47044), *Klebsiella pneumoniae* (B02624), *Pseudomonas aeruginosa* (G445), *Enterobacter cloacae* (R1288), *Proteus mirabilis* (G464), *Klebsiella aerogenes* (DSM 30053) and *Aeromonas hydrophila* (DSM 30187).
- Confirmation tests – Julian et al. approach 2 wells from each of 60 QTs. Key recommendation is to use mTEC rather than MacConkey in future.
- Deep 96 well plates used, as well as IDEXX Legiolert Quantitrays
- Analysis: P/A sensitivity and specificity vs reference method; quantitative *E. coli* risk levels concordance and *t*-test for mean difference in counts

Bain et al. 2015. Evaluation of an Inexpensive Growth Medium for Direct Detection of Escherichia coli in Temperate and Sub-Tropical Waters. https://doi.org/10.1371/journal.pone.0140997

- Reference methods: Colilert QT, EC-MUG (EC-MUG/MacConkey and API 20E used for confirmations)
- Control strains: Control strains were *E. coli* (NCTC 9001), non-*E. coli* total coliform (*Klebsiella pneumoniae* NCTC 9633) and non-coliform (*Pseudomonas aeruginosa* NCTC 10662)
- ~200 samples of water tested vs Colilert (14 sources in UK, 25 source in SA)
- Confirmatory testing of 400-600 wells vs EC-MUG then API 20E if inconsistent
- Analysis: P/A: confirmatory test sensitivity and specificity, Quantitative results: difference vs mean, Spearman's rank