Australian Government
Department of Industry, Science and Resources

National Measurement Institute

Proficiency Test Final Report
AQA 24-19
Hydrocarbons in River Water

March 2025

© Commonwealth of Australia 2025.

Unless otherwise noted, the Commonwealth owns the copyright (and any other intellectual property rights, if any) in this publication.

All material in this publication is provided under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0), with the exception of:

- the Commonwealth Coat of Arms;
- the logo of the Department of Industry, Science and Resources;
- photographs of our staff and premises; and
- content supplied by third parties.

Creative Commons Attribution 4.0 International Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. A summary of the licence terms is available at: creativecommons.org/licenses/by/4.0/. Further details are available on the Creative Commons website, at: creativecommons.org/licenses/by/4.0/legalcode.

You may not copy, distribute, transmit or adapt any material in this publication in any way that suggests that this department or the Commonwealth endorses you or any of your services or products.

Attribution

Material contained in this publication is to be attributed to this department as:

© Commonwealth of Australia, Department of Industry, Science and Resources, Proficiency Test Final Report AQA 24-19 Hydrocarbons in River Water, 2025.

Third party copyright

Wherever a third party holds copyright in material contained in this publication, the copyright remains with that party. Their permission may be required to use the material.

This department has made all reasonable efforts to:

- clearly label material where the copyright is owned by a third party;
- ensure that the copyright owner has consented to this material being contained in this publication.

Using the Commonwealth Coat of Arms

The terms of use for the Coat of Arms are available on the Department of Prime Minister and Cabinet's website, at www.pmc.gov.au/resource-centre/government/commonwealth-coat-arms-information-and-guidelines

ACKNOWLEDGMENTS

This study was conducted by the National Measurement Institute Australia (NMIA). Support funding was provided by the Australian Government Department of Industry, Science and Resources.

I would like to thank the management and staff of the participating laboratories for supporting the study. It is only through widespread participation that we can provide an effective service to laboratories.

The assistance of the following NMIA staff members in the planning, conduct and reporting of the study is acknowledged.

Jenny Xu

Geoff Morschel

Jasmine Duong

Aaron Mamo

Lesley Johnston

Raluca Iavetz

Manager, Chemical Reference Values

105 Delhi Road, North Ryde, NSW 2113, Australia

Phone: +61 2 9449 0178

Email: raluca.iavetz@measurement.gov.au

NATA WORLD RECOGNISED ACCREDITATION

Accredited for compliance with ISO/IEC 17043

TABLE OF CONTENTS

THIS PAGE IS INTENTIONALLY BLANK

SUMMARY

AQA 24-19 Hydrocarbons in River Water commenced in September 2024. Fifteen laboratories registered to participate, and 14 participants submitted results.

The sample set consisted of three water samples. Samples were prepared in the Sydney NMIA laboratory using river water. Participants measured total recoverable hydrocarbons (TRH) in Sample S1, volatile hydrocarbons (C6 to C10), benzene, toluene, ethylbenzene and xylenes (BTEX) in Sample S2, and polycyclic aromatic hydrocarbons (PAHs) in Sample S3.

The assigned values for all scored analytes were the robust averages of participants' results. The associated uncertainties were evaluated from the robust standard deviations of participants' results.

Traceability: The consensus of participants' results is not traceable to any external reference, so although expressed in SI units, metrological traceability has not been established.

The outcomes of the study were assessed against the aims as follows:

- *Compare the performances of participants and assess their accuracy in the identification and measurement of hydrocarbons in river water.*

Laboratories **1**, **2**, **3**, **4**, **6**, **7**, **8**, **9**, **10**, **11**, **12**, **14** and **15** reported results for all 11 scored analytes.

One participant (Laboratory **14**) did not report a numeric result for an analyte they tested for and was present in the sample (total of one result). Three participants reported numeric results for analytes that were not spiked into the samples (total of nine results).

Of 148 z-scores, 125 (84%) returned a score of $|z| \leq 2.0$, indicating an acceptable performance.

Of 143 E_n-scores, 101 (71%) returned a score of $|E_n| < 1.0$, indicating agreement of the participant's result with the assigned value within their respective expanded uncertainties.

Laboratories **6**, **7**, **9** and **14** returned acceptable z-scores and E_n-scores for all 11 scored analytes.

- *Develop the practical application of traceability and measurement uncertainty, and provide participants with information that will be useful in assessing their uncertainty evaluations.*

Of 222 numeric results, 167 (75%) were reported with an associated expanded measurement uncertainty. Reported expanded uncertainties were within the range of 1.9% to 67% relative.

- *Evaluate the laboratories' test methods.*

For TRH analysis, the most common methodology was liquid-liquid extraction with dichloromethane, and analysis using GC-FID. Three participants used solid phase extraction with hexane instead, and these participants reported significantly higher TRH results.

For BTEX analysis, the most common methodology was purge-and-trap GC-MS.

For PAHs analysis, the most common methodology was liquid-liquid extraction with dichloromethane, and analysis using GC-MS or GC-MS/MS.

For both BTEX and PAHs analysis, results reported by participants were generally compatible with each other, regardless of the method used.

- *Produce materials that can be used in method validation and as control samples.*

The test samples of this proficiency study are homogeneous and are well characterised. Surplus samples are available for purchase from NMIA and can be used for quality control and method validation purposes.

1 INTRODUCTION

1.1 NMIA Proficiency Testing Program

The National Measurement Institute Australia (NMIA) is responsible for Australia's national measurement infrastructure, providing a range of services including a chemical proficiency testing program.

Proficiency testing (PT) is the 'evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons'.[1] NMIA PT studies target chemical testing in areas of high public significance such as trade, environment, law enforcement and food safety. NMIA offers studies in:

- pesticide residues in fruit, vegetables and herbs, water and soil;
- petroleum hydrocarbons in water and soil;
- per- and polyfluoroalkyl substances in water, soil, biota and food;
- inorganic analytes in water, soil, filters, food and pharmaceuticals;
- controlled drug assay, drugs in wipes and clandestine laboratory; and
- allergens in food.

1.2 Study Aims

The aims of the study were to:

- compare the performances of participants and assess their accuracy in the identification and measurement of hydrocarbons in river water;
- develop the practical application of traceability and measurement uncertainty, and provide participants with information that will be useful in assessing their uncertainty evaluations;
- evaluate the laboratories' test methods; and
- produce materials that can be used in method validation and as control samples.

The choice of the test method was left to the participating laboratories.

1.3 Study Conduct

The conduct of NMIA proficiency tests is described in the NMIA Study Protocol for Proficiency Testing.[2] The statistical methods used are described in the NMIA Chemical Proficiency Testing Statistical Manual.[3] These documents have been prepared with reference to ISO/IEC 17043,[1] and The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories.[4]

NMIA is accredited by the National Association of Testing Authorities, Australia (NATA) to ISO/IEC 17043:2023 as a provider of proficiency testing schemes.[1] This study falls within the scope of NMIA's accreditation.

2 STUDY INFORMATION

2.1 Study Timetable

The timetable of this study was:

Invitations sent	30/9/2024
Samples sent	21/10/2024
Results due	15/11/2024
Interim Report	18/11/2024
Preliminary Report	21/11/2024

2.2 Participation

Fifteen laboratories registered to participate, and all participants were assigned a confidential laboratory code number for this study. Fourteen participants submitted results.

2.3 Selection of Analytes

The analytes and their concentrations in this study were typical of those encountered by environmental testing laboratories monitoring river water to assess the impact of transport fuels in the environment, or the contamination from industry that entails the use of wood, petroleum or coal to generate heat and power. Investigation levels for the analytes are set out in the National Environmental Protection (Assessment of Site Contamination) Measure, Schedule B1 *Guideline on Investigation Levels for Soil and Groundwater*.[5]

For Sample S1, participants were requested to measure semi-volatile hydrocarbons (>C10-C40) and total recoverable hydrocarbons (TRH). For Sample S2, participants were requested to measure volatile hydrocarbons (C6-C10), and benzene, toluene, ethylbenzene and xylenes (BTEX). Participants were provided with a list of potential poly-aromatic hydrocarbons (PAHs) spiked into Sample S3 (Table 1).

Table 1 Possible Spiked PAHs in Sample S3

Naphthalene	Fluorene	Benz[*a*]anthracene	Benzo[*a*]pyrene
Acenaphthylene	Phenanthrene	Chrysene	Indeno[1,2,3-*cd*]pyrene
Acenaphthene	Fluoranthene	Benzo[*b*]fluoranthene	Dibenz[*a,h*]anthracene
Anthracene	Pyrene	Benzo[*k*]fluoranthene	Benzo[*g,h,i*]perylene

2.4 Test Material Preparation

Three test samples were prepared by spiking river water with various analytes to obtain the concentrations listed in Table 2. Sample S1 was spiked with diesel fuel, Sample S2 was spiked with unleaded petrol and diesel fuel, and Sample S3 was spiked with different amounts of acenaphthene, anthracene, benz[*a*]anthracene, benzo[*a*]pyrene, chrysene, fluoranthene fluorene and pyrene.

Table 2 Spiked Values of Test Samples

Sample	Analyte	Spiked Value (µg/L)	Uncertainty* (µg/L)
S1**	TRH	1900	100
S2	Benzene	55.1	2.8
	Toluene	128	6
	Ethylbenzene	24.4	1.2

Sample	Analyte	Spiked Value (µg/L)	Uncertainty* (µg/L)
	Xylenes	109	5
	Total BTEX	316	16
S3	Acenaphthene	15.9	0.8
	Anthracene	4.48	0.22
	Benz[*a*]anthracene	4.50	0.22
	Benzo[*a*]pyrene	2.00	0.10
	Chrysene	3.01	0.15
	Fluoranthene	3.00	0.15
	Fluorene	3.01	0.15
	Pyrene	2.00	0.10

* Evaluated expanded uncertainty at approximately 95% confidence using a coverage factor of 2. Stability was not considered and so the expanded uncertainty is related to the concentration at the time of spiking.

** Sample S1 was spiked with diesel. Spiked value is approximate.

Additional information on sample preparation is given in Appendix 1.

2.5 Homogeneity and Stability of Test Materials

No homogeneity or stability testing was conducted on these test materials before the samples were sent. The samples were prepared, stored and dispatched using a process that has been demonstrated to produce sufficiently homogeneous and stable samples in previous similar NMIA PT studies. Additionally, the storage stability of petroleum hydrocarbons in water has been previously established.[6]

Participants' results also gave no reason to question the homogeneity or stability of Samples S2 and S3. Analytes have only been scored if there was a reasonable consensus between participants' results, and if the consensus value to spiked value ratio was similar to those observed in previous NMIA studies. There was a higher variation between participants' results observed for Sample S1, and no assigned values could be set for the analytes in this sample. Three bottles of this sample were analysed to conduct partial homogeneity testing, and no evidence of inhomogeneity was detected.

Additional information is given in Appendix 2.

2.6 Test Material Storage, Dispatch and Receipt

The test samples were stored at approximately 4 °C prior to dispatch. Samples were dispatched on 21 October 2024.

The following items were also sent to participants:

- a covering letter which included a description of the test samples and instructions for participants; and
- a form for participants to confirm the receipt and condition of the test samples.

An Excel spreadsheet for the electronic reporting of results was emailed to participants.

2.7 Instructions to Participants

Participants were instructed as follows:

- Quantitatively analyse the samples using your routine test method.

- Do not test for volatile hydrocarbons (C6-C10) or BTEX components in Sample S1.
- Participants need not test for all listed analytes.
- If analyses cannot be commenced on the day of receipt, please store the samples chilled.
- For each analyte in each sample, report a single result in units of µg/L expressed as if reporting to a client, applying the limit of reporting of the method used for analysis. This is the figure that will be used in all statistical analysis in the study report.
- For each analyte in each sample, report the associated expanded uncertainty in units of µg/L (e.g. 2000 ± 200 µg/L), if determined.
- Report results for the following:
    - Sample S1: Semi-volatile hydrocarbons (>C10-C40) and Total Recoverable Hydrocarbons (TRH). Use your laboratory's chosen quantitation range, and indicate what this range is. Australian NEPM fractions >C10-C16, >C16-C34, >C34-C40 are encouraged. The concentration range is between 200 – 10000 µg/L.
    - Sample S2: Volatile Hydrocarbons (C6-C10), Benzene, Toluene, Ethylbenzene, Total Xylenes and Total BTEX. Individual BTEX components concentration is between 10 – 2500 µg/L.
    - Sample S3: PAHs. The concentration range is between 1 – 50 µg/L.
- Give details of your methodology and basis of uncertainty as requested by the results sheet emailed to you.
- Return the completed results sheet by 15 November 2024 by email to proficiency@measurement.gov.au.

2.8 Interim Report and Preliminary Report

An Interim Report was emailed to all participants on 18 November 2024.

A Preliminary Report was emailed to all participants on 21 November 2024. This report included a summary of the results reported by participants, assigned values, performance coefficients of variation (PCVs), *z*-scores and E_n-scores for each analyte in this study. No scores or statistics from the Preliminary Report have been changed in the present Final Report. Results from Laboratory **13** in Sample S3 have been modified from NT to NS, to reflect that they were not supplied this sample.

3 PARTICIPANT LABORATORY INFORMATION

3.1 Participants' Test Methods

Participants were requested to provide information about their test methods. Responses are presented in Appendix 4.

3.2 Basis of Participants' Measurement Uncertainty Evaluation

Participants were requested to provide information about their basis of measurement uncertainty (MU). Responses are presented in Table 3. Some responses may be modified so that the participant cannot be identified. Participants reported using the same basis of uncertainty evaluation across all samples analysed.

Table 3 Basis of Uncertainty Evaluation

Lab. Code	Approach to Evaluating MU	Information Sources for MU Evaluation*		Guide Document for Evaluating MU
		Precision	Method Bias	
1	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - SS Duplicate analysis Instrument calibration	Laboratory bias from PT studies Instrument calibration Recoveries of SS Standard purity	Eurachem/CITAC Guide
2	Bottom Up (ISO/GUM, fish bone/cause and effect diagram) Coverage factor not reported	Instrument calibration	Instrument calibration Recoveries of SS Standard purity	ISO/GUM
3	Standard deviation of replicate analyses multiplied by 2 or 3 k = 2	Standard deviation from PT studies only		ISO/GUM
4	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - SS		Eurachem/CITAC Guide
6	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - CRM Duplicate analysis Instrument calibration	CRM Instrument calibration Recoveries of SS	NMI Uncertainty Course
7	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Standard deviation from PT studies only		ISO/GUM
8	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - RM Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS	Eurachem/CITAC Guide
9	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS	Recoveries of SS	ISO/GUM

Lab. Code	Approach to Evaluating MU	Information Sources for MU Evaluation*		Guide Document for Evaluating MU
		Precision	Method Bias	
	Coverage factor not reported			
10	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - SS	Recoveries of SS	ISO/GUM
11	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - RM Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS	Eurachem/CITAC Guide
12	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - SS Duplicate analysis	Instrument calibration Recoveries of SS Standard purity	
13	Standard deviation of replicate analyses multiplied by 2 or 3 k = 2	Control samples - SS Duplicate analysis	Recoveries of SS	Eurachem/CITAC Guide
14	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - SS Duplicate analysis	Instrument calibration Recoveries of SS	
15	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - RM Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS	Eurachem/CITAC Guide

* CRM = Certified Reference Material; RM = Reference Material; SS = Spiked Samples

3.3 Participants' Comments

Participants were invited to make any comments or suggestions on the samples, this study, or possible future studies. Such feedback may be useful in improving future studies. Participants' comments are presented in Table 4. Some responses may be modified so that the participant cannot be identified.

Table 4 Participants' Comments

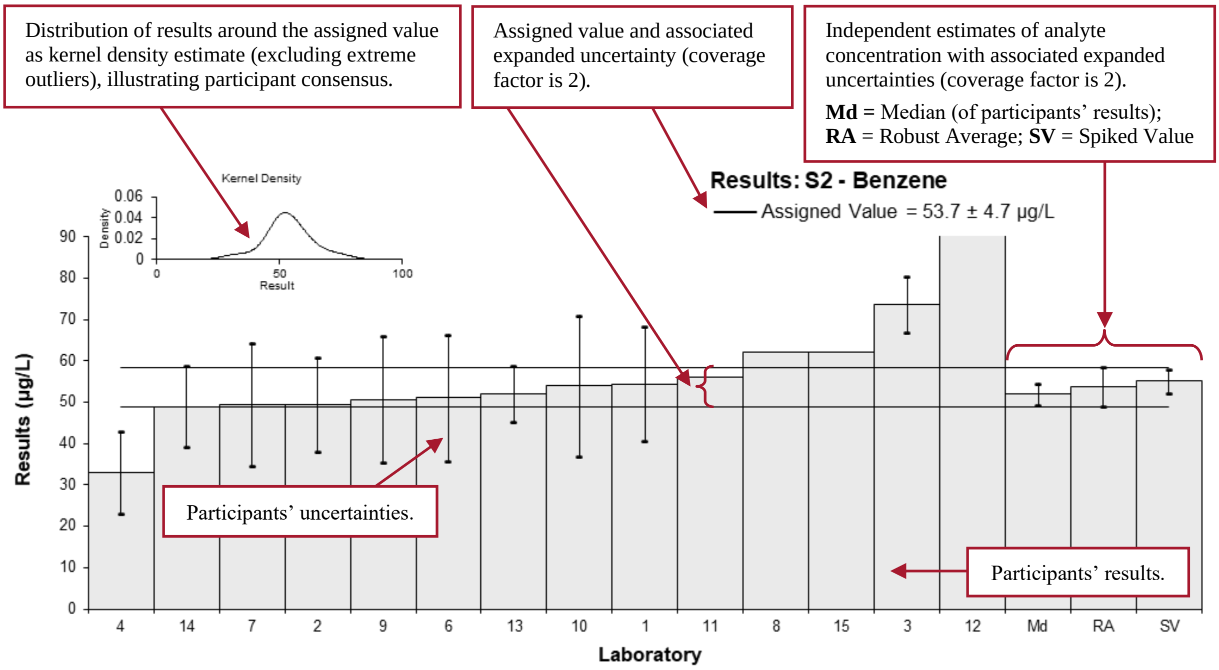
Lab. Code	Sample	Participant's Comments	Study Coordinator's Response
8	S1	Noted variation between analysts Methodology: Diluted	
	All	Uncertainty: MU omitted as currently updating	
11	S1	Noted variation between analysts Methodology: Diluted	
	S3	Naphthalene variation	
	All	Uncertainty: MU omitted as currently updating	

Lab. Code	Sample	Participant's Comments	Study Coordinator's Response
13	All	Have a drop down for 'Lab. Code' that has the participating labs listed.	As per ISO/IEC 17043 requirements, participants in our PT studies are confidential. Therefore, we cannot include a list of participating labs in the results sheet as this would reveal their identities.
15	S1	Noted variation between analysts Methodology: Diluted	
	All	Uncertainty: MU omitted as currently updating	

4 PRESENTATION OF RESULTS AND STATISTICAL ANALYSIS

4.1 Results Summary

Participant results are listed in Tables 5 to 23 with summary statistics: robust average, median, mean, number of numeric results (N), maximum (Max), minimum (Min), robust standard deviation (Robust SD) and robust coefficient of variation (Robust CV), along with other estimates of analyte concentration. Bar charts of results and performance scores are presented in Figures 2 to 19. An example chart with interpretation guide is shown in Figure 1.



Distribution of results around the assigned value as kernel density estimate (excluding extreme outliers), illustrating participant consensus.

Assigned value and associated expanded uncertainty (coverage factor is 2).

Independent estimates of analyte concentration with associated expanded uncertainties (coverage factor is 2).
Md = Median (of participants' results);
RA = Robust Average; **SV** = Spiked Value

Participants' uncertainties.

Participants' results.

Figure 1 Guide to Presentation of Results

4.2 Outliers, Extreme Outliers and Other Excluded Results

Outliers were results less than 50% and greater than 150% of the robust average, and these were removed before the calculation of the assigned value.[3,4] Extreme outliers were obvious blunders, e.g. results reported with incorrect units or basis, and such results were removed for the calculation of all summary statistics.[3,4]

The results from Laboratory **1** in Sample S3 were consistently lower than the consensus of participants' results by approximately the same factor. The results reported by Laboratory **3** for Sample S3 were most likely for a different sample, as the results reported for spiked analytes were inconsistent with the consensus of participants' results, and numeric results were also reported for many analytes not spiked into this sample. To avoid unfair scoring from a bias in the assigned value, these results were excluded from the robust average calculations; they were also excluded from the calculation of all summary statistics.

4.3 Assigned Value

The assigned value is defined as the 'value attributed to a particular property or characteristic of a proficiency testing item'.[1] In this PT study, the property is the concentration of the analytes in the samples. Assigned values were the robust averages of participants' results, and the expanded uncertainties were evaluated from the associated robust SDs (Appendix 3).

4.4 Robust Average and Robust Between-Laboratory Coefficient of Variation

The robust averages and associated expanded MUs, and robust between-laboratory CVs (a measure of the variability of participants' results) were calculated using the procedure described in ISO 13528.[7]

4.5 Performance Coefficient of Variation

The performance coefficient of variation (PCV) is a fixed measure of the between-laboratory variation that in the judgement of the study coordinator would be expected from participants, given the analyte concentrations. The PCV is not the CV of participants' results; it is set by the study coordinator and is based on the concentration of the analytes and experience from previous studies, and is supported by mathematical models such as the Thompson-Horwitz equation.[8] By setting a fixed and realistic value for the PCV, a participant's performance does not depend on other participants' performances and can be compared from study to study.

4.6 Target Standard Deviation for Proficiency Assessment

The target standard deviation for proficiency assessment (σ) is the product of the assigned value (X) and the PCV, as presented in Equation 1.

$$\sigma = X \times PCV \qquad \textit{Equation 1}$$

4.7 *z*-Score

For each participant result, a *z*-score is calculated according to Equation 2.

$$z = \frac{(\chi - X)}{\sigma} \qquad \textit{Equation 2}$$

where:

z is *z*-score

χ is a participant's result

X is the assigned value

σ is the target standard deviation for proficiency assessment from Equation 1

For the absolute value of a *z*-score ($|z|$):

- $|z| \leq 2.0$ is acceptable;
- $2.0 < |z| < 3.0$ is questionable; and
- $|z| \geq 3.0$ is unacceptable.

To account for potential low bias in consensus value due to inefficient methodologies, scores may be adjusted for a 'maximum acceptable result'. Additional information is given in Section 6.3.

4.8 *E_n*-Score

The E_n-score is complementary to the *z*-score in assessment of laboratory performance. E_n-score includes measurement uncertainty and is calculated according to Equation 3.

$$E_n = \frac{(\chi - X)}{\sqrt{{U_\chi}^2 + {U_X}^2}} \qquad \textit{Equation 3}$$

where:

E_n is E_n-score

χ is a participant's result

X is the assigned value

U_χ is the expanded uncertainty of the participant's result

U_X is the expanded uncertainty of the assigned value

For the absolute value of an E_n-score ($|E_n|$):

- $|E_n| < 1.0$ is acceptable; and
- $|E_n| \geq 1.0$ is unacceptable.

4.9 Traceability and Measurement Uncertainty

Laboratories accredited to ISO/IEC 17025 must establish and demonstrate the traceability and measurement uncertainty associated with their test results.[9]

Guidelines for quantifying uncertainty in analytical measurement are described in the Eurachem/CITAC Guide.[10]

5 TABLES AND FIGURES

Table 5

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	>C10-C16
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1	570	166
2	648	156
3	1454	320
4	NR	NR
6	320	96
7	330	99
8	302	NR
9	202.4	80.96
10	350	106
11	2788	NR
12	65	19.5
13	NR	NR
14	<25	NR
15	4279	NR

Statistics

Assigned Value	Not Set	
Robust Average	730	580
Median	350	250
Mean	1030	
N	11	
Max	4279	
Min	65	
Robust SD	760	
Robust CV	110%	

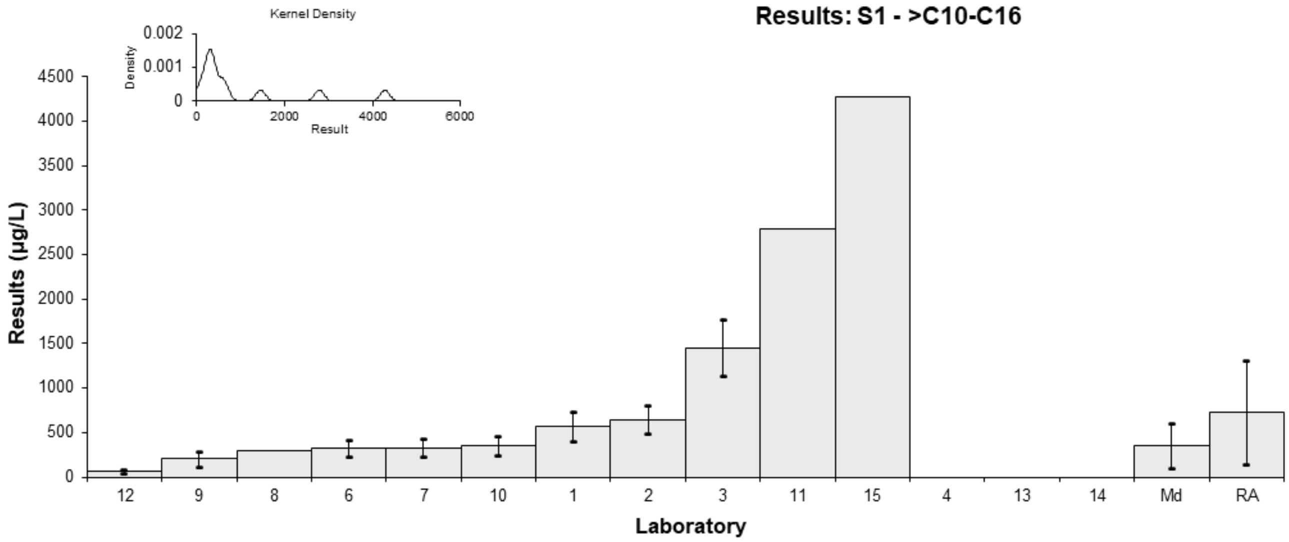

Kernel Density
0.002
0.001
0
Density
0
2000
4000
6000
Result
Results: S1 - >C10-C16
4500
4000
3500
3000
2500
2000
1500
1000
500
0
Results (µg/L)
12
9
8
6
7
10
1
2
3
11
15
4
13
14
Md
RA
Laboratory

Figure 2

Table 6

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	>C16-C34
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1	420	122
2	557	161
3	673	225
4	NR	NR
6	360	108
7	372	111.6
8	<100	NR
9	304.2	121.68
10	371	113
11**	7224	NR
12	68	20.4
13	NR	NR
14	<100	NR
15	578	NR

** Extreme Outlier, see Section 4.2

Statistics

Assigned Value	Not Set	
Robust Average	420	150
Median	372	84
Mean	410	
N	9	
Max	673	
Min	68	
Robust SD	180	
Robust CV	43%	

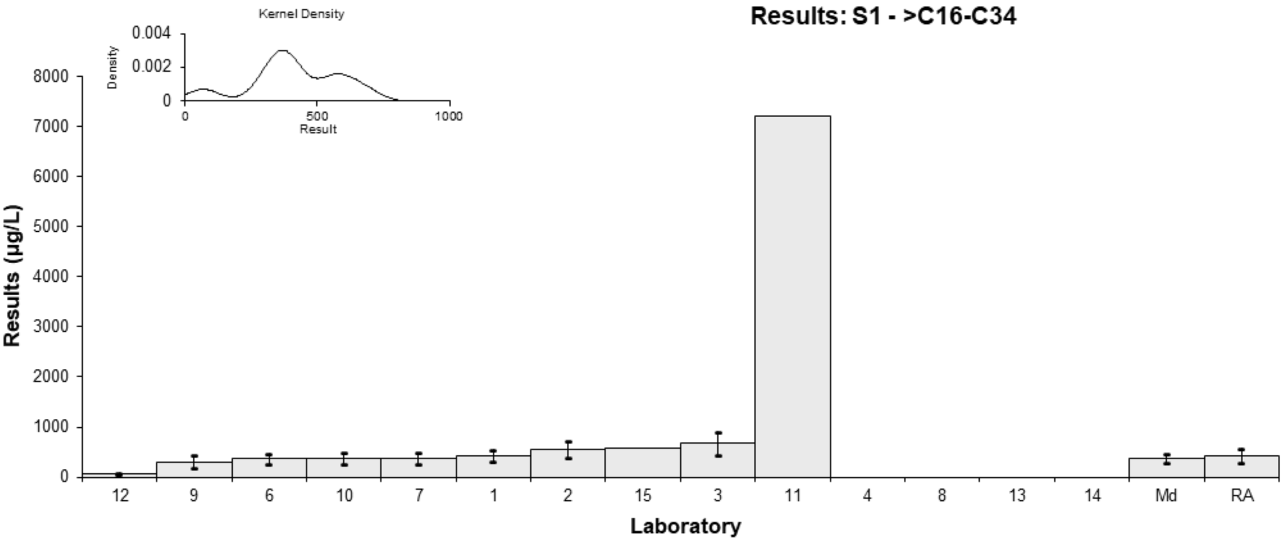

Results: S1 - >C16-C34
Kernel Density
Density
0.004
0.002
0
0
500
1000
Result
Results (µg/L)
8000
7000
6000
5000
4000
3000
2000
1000
0
12
9
6
10
7
1
2
15
3
11
4
8
13
14
Md
RA
Laboratory

Figure 3

Table 7

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	>C34-C40
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1	20	6
2	<200	NR
3	< 100	< 100
4	NR	NR
6	< 100	NR
7	< 100	30
8	5522	NR
9	<100	NR
10	< 100	30
11	1466	NR
12	<50	NR
13	NR	NR
14	<100	NR
15	627	NR

Statistics

Assigned Value	Not Set	
Robust Average	NA (N<6)	
Median	1000	
Mean	1900	
N	4	
Max	5522	
Min	20	
Robust SD	NA (N<6)	
Robust CV	NA (N<6)	

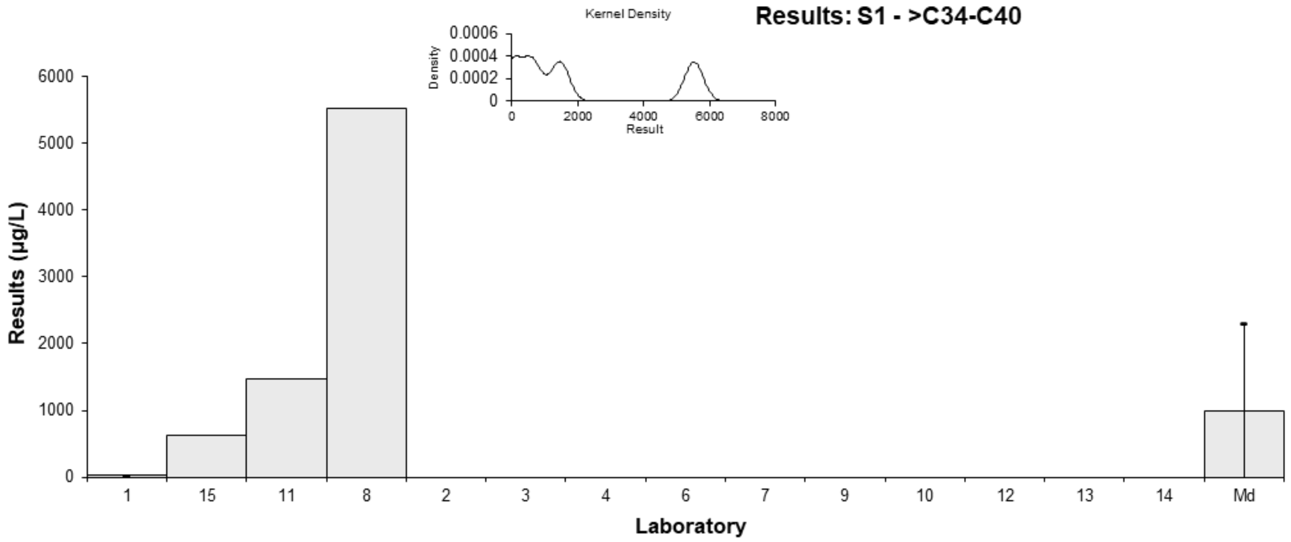

Kernel Density
Density
0.0006
0.0004
0.0002
0
0
2000
4000
6000
8000
Result
Results: S1 - >C34-C40
Results (µg/L)
6000
5000
4000
3000
2000
1000
0
1
15
11
8
2
3
4
6
7
9
10
12
13
14
Md
Laboratory

Figure 4

Table 8

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	TRH
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1	1010	293
2	1205	NR
3	2126	545
4	<500	NR
6	680	204
7	702	240.6
8	5824	NR
9	506.6	202.64
10	721	219
11	11479	NR
12	133	39.9
13	1597	587
14	<250	NR
15	5484	NR

Statistics

Assigned Value	Not Set	
Spike Value	1900	100
Robust Average	2200	1800
Median	1110	580
Mean	2600	
N	12	
Max	11479	
Min	133	
Robust SD	2500	
Robust CV	120%	

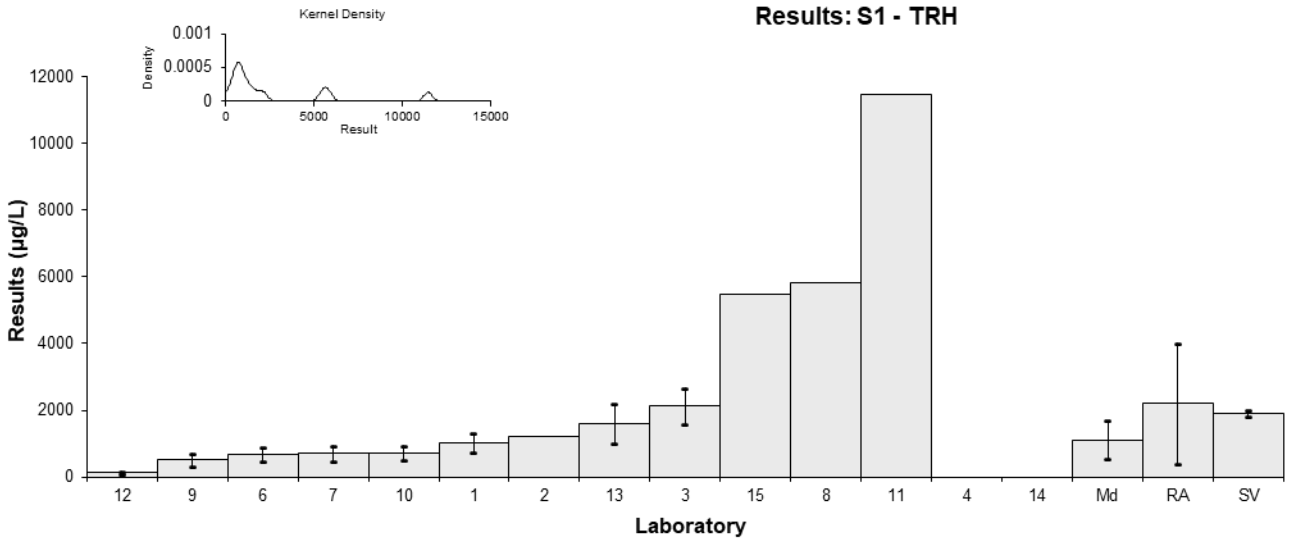

Kernel Density
0.001
0.0005
0
Density
0
5000
10000
15000
Result
Results: S1 - TRH
Results (µg/L)
12000
10000
8000
6000
4000
2000
0
12
9
6
7
10
1
2
13
3
15
8
11
4
14
Md
RA
SV
Laboratory

Figure 5

Table 9 Non-NEPM Hydrocarbon Ranges Reported by Participants for Sample S1

Lab. Code	Range	Result (µg/L)	Uncertainty (µg/L)
4	C7-C9	<200	NR
	C10-C14	200	60
	C15-C36	<300	NR
13	C10-C14	500	170
	C15-C28	1097	417
	C29-C36	<80	NR

THIS PAGE IS INTENTIONALLY BLANK

Table 10

Sample Details

Sample No.	S2
Matrix	River Water
Analyte	C6-C10
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1	NT	NT
2	548.48	67.46
3	155.5	6.5
4	NT	NT
6	745.5	186
7	760.2	228.06
8	447	NR
9	592	177.6
10	629	189
11	398	NR
12	53	15.9
13	NT	NT
14	500	100
15	537	NR

Statistics

Assigned Value	Not Set	
Robust Average	500	170
Median	540	100
Mean	490	
N	11	
Max	760.2	
Min	53	
Robust SD	230	
Robust CV	46%	

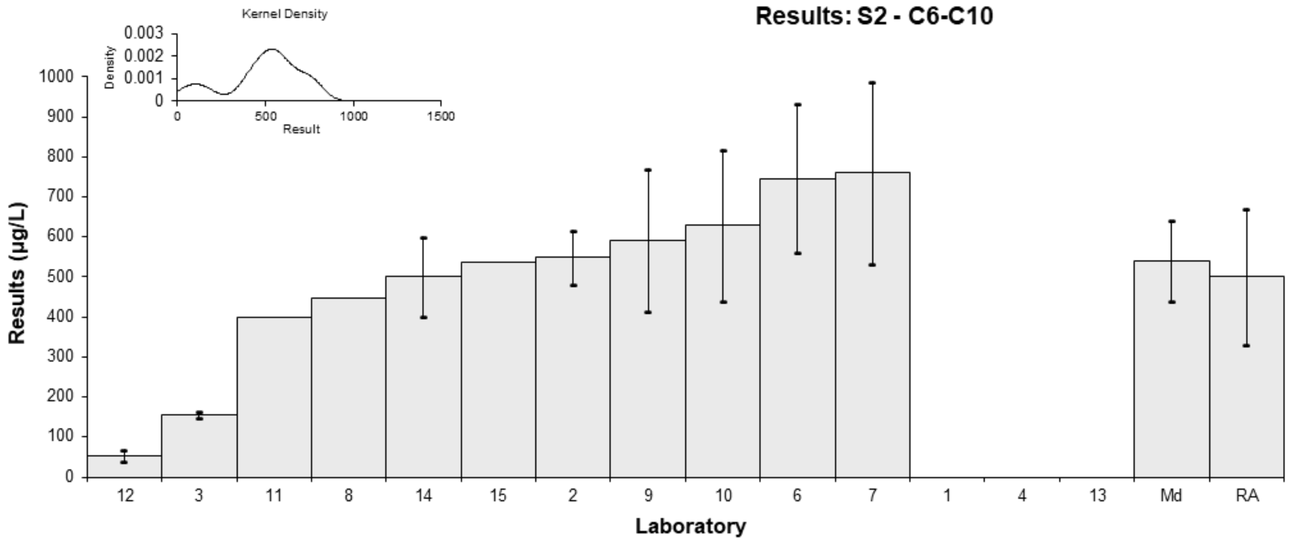

Kernel Density
Density
0.003
0.002
0.001
0
0
500
1000
1500
Result
Results: S2 - C6-C10
Results (µg/L)
1000
900
800
700
600
500
400
300
200
100
0
12
3
11
8
14
15
2
9
10
6
7
1
4
13
Md
RA
Laboratory

Figure 6

Table 11

Sample Details

Sample No.	S2
Matrix	River Water
Analyte	Benzene
Unit	µg/L

Participant Results

Lab. Code	**Result**	**Uncertainty**	**z**	**E_n**
1	54.5	13.8	0.10	0.05
2	49.55	11.4	-0.52	-0.34
3	73.7	6.7	2.48	2.44
4	33	9.9	-2.57	-1.89
6	51.1	15.3	-0.32	-0.16
7	49.5	14.85	-0.52	-0.27
8	62	NR	1.03	1.77
9	50.7	15.21	-0.37	-0.19
10	54	17	0.04	0.02
11	56	NR	0.29	0.49
12**	764	229.2	88.18	3.10
13	52	6.8	-0.21	-0.21
14	49	9.8	-0.58	-0.43
15	62	NR	1.03	1.77

** Extreme Outlier, see Section 4.2

Statistics

Assigned Value	53.7	4.7
Spike Value	55.1	2.8
Robust Average	53.7	4.7
Median	52.0	2.6
Mean	53.6	
N	13	
Max	73.7	
Min	33	
Robust SD	6.7	
Robust CV	13%	

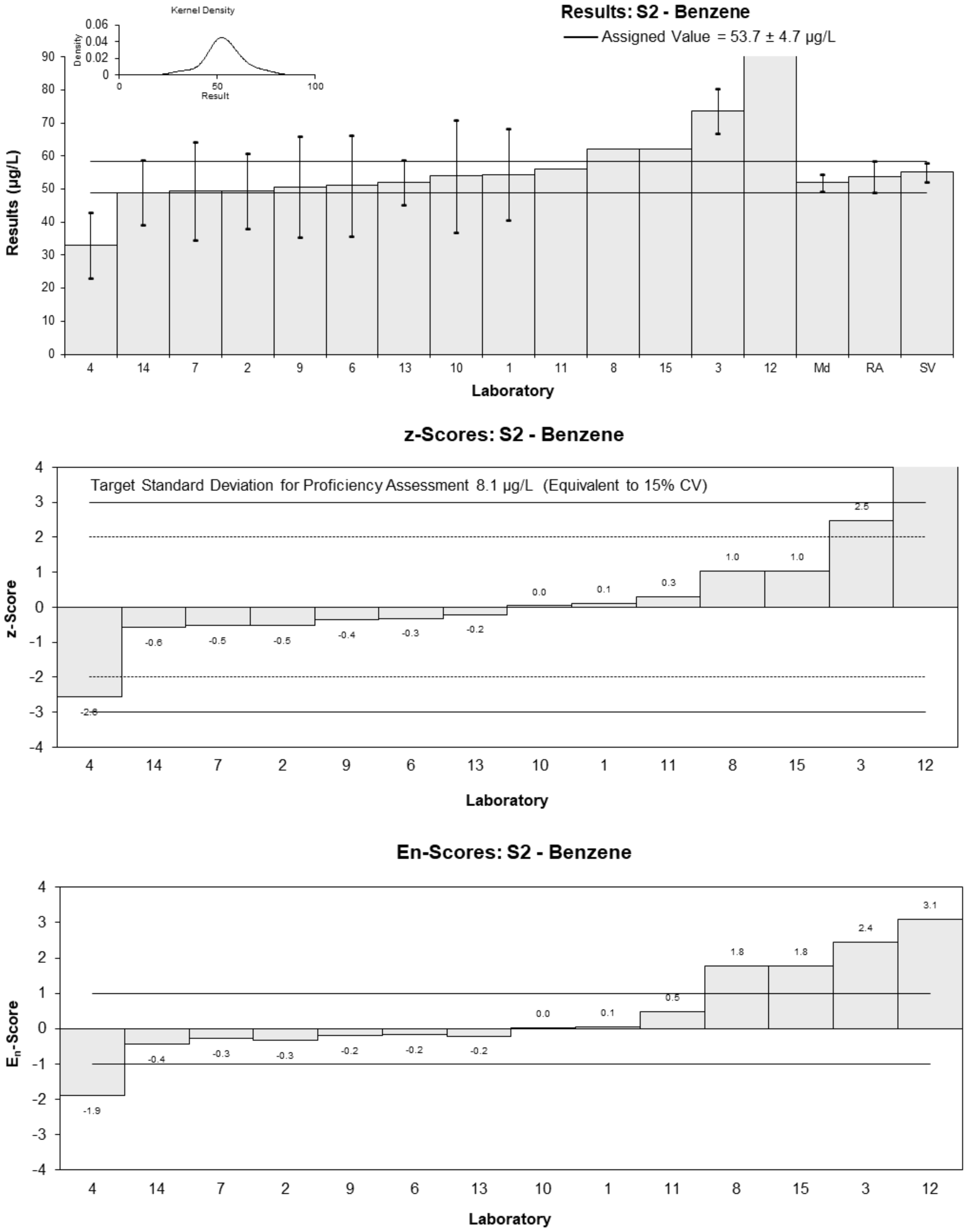

Kernel Density
Density
0.06
0.04
0.02
0
0
50
100
Result
Results: S2 - Benzene
Assigned Value = 53.7 ± 4.7 µg/L
Results (µg/L)
90
80
70
60
50
40
30
20
10
0
4
14
7
2
9
6
13
10
1
11
8
15
3
12
Md
RA
SV
Laboratory
z-Scores: S2 - Benzene
Target Standard Deviation for Proficiency Assessment 8.1 µg/L (Equivalent to 15% CV)
z-Score
4
3
2
1
0
-1
-2
-3
-4
-2.6
-0.6
-0.5
-0.5
-0.4
-0.3
-0.2
0.0
0.1
0.3
1.0
1.0
2.5
4
14
7
2
9
6
13
10
1
11
8
15
3
12
Laboratory
En-Scores: S2 - Benzene
En-Score
4
3
2
1
0
-1
-2
-3
-4
-1.9
-0.4
-0.3
-0.3
-0.2
-0.2
-0.2
0.0
0.1
0.5
1.8
1.8
2.4
3.1
4
14
7
2
9
6
13
10
1
11
8
15
3
12
Laboratory

Figure 7

Table 12

Sample Details

Sample No.	S2
Matrix	River Water
Analyte	Toluene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	93.5	19.2	-0.73	-0.51
2	94.71	21.78	-0.65	-0.41
3	135.6	5.6	1.94	2.31
4	65.5	19.65	-2.51	-1.72
6	109.4	27.4	0.28	0.15
7	105.3	31.59	0.02	0.01
8	129	NR	1.52	2.00
9	95.5	28.65	-0.60	-0.31
10	114	33	0.57	0.26
11	110	NR	0.32	0.42
12*	633	189.9	33.52	2.77
13	97	19.4	-0.51	-0.35
14	92	18	-0.83	-0.60
15	116	NR	0.70	0.92

* Outlier, see Section 4.2

Statistics

Assigned Value	105	12
Spike Value	128	6
Robust Average	108	13
Median	107	12
Mean	142	
N	14	
Max	633	
Min	65.5	
Robust SD	20	
Robust CV	19%	

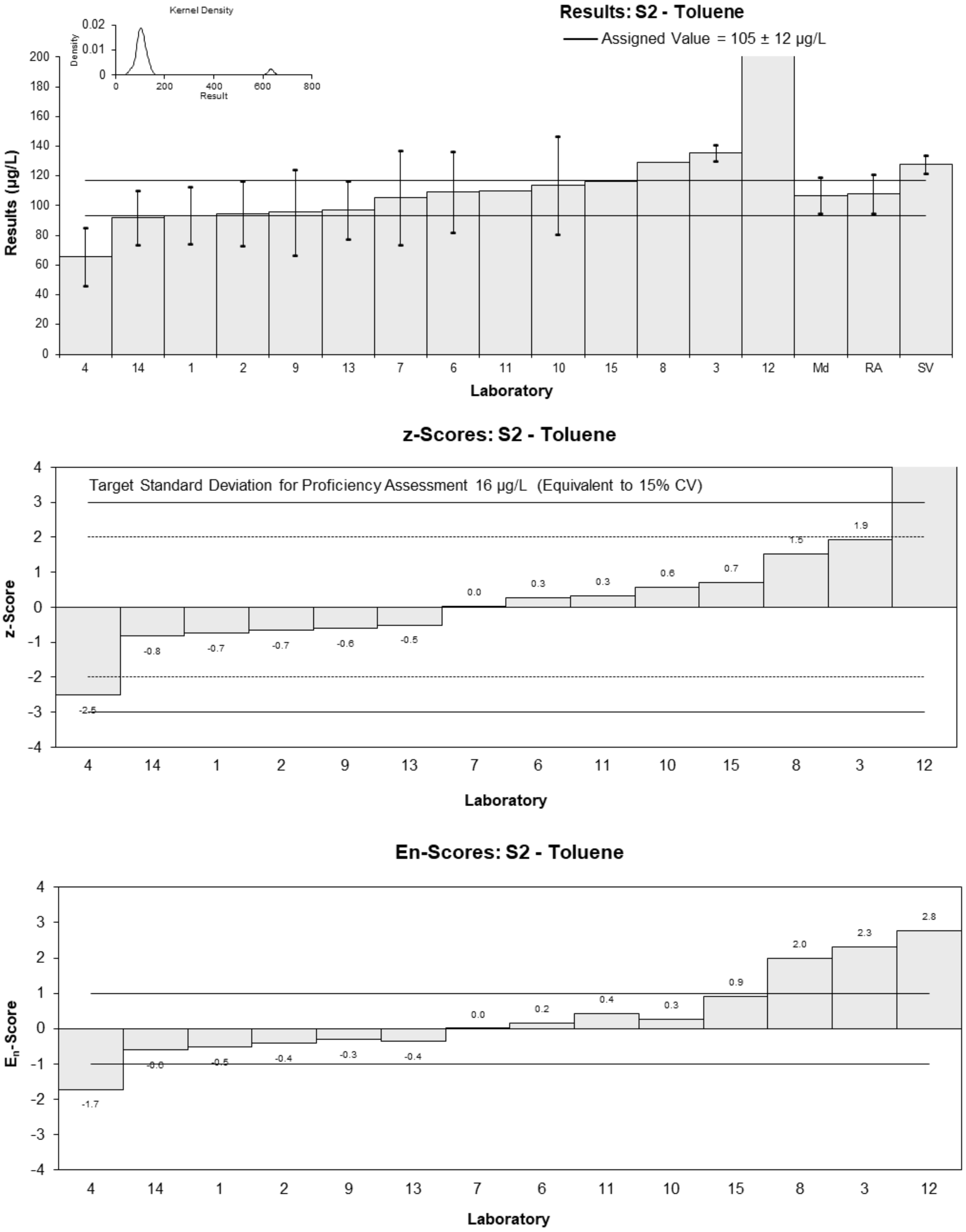

Results: S2 - Toluene
Assigned Value = 105 ± 12 µg/L
Kernel Density
Density
0.02
0.01
0
0
200
400
600
800
Result
Results (µg/L)
200
180
160
140
120
100
80
60
40
20
0
4
14
1
2
9
13
7
6
11
10
15
8
3
12
Md
RA
SV
Laboratory
z-Scores: S2 - Toluene
Target Standard Deviation for Proficiency Assessment 16 µg/L (Equivalent to 15% CV)
z-Score
4
3
2
1
0
-1
-2
-3
-4
-2.5
-0.8
-0.7
-0.7
-0.6
-0.5
0.0
0.3
0.3
0.6
0.7
1.5
1.9
4
14
1
2
9
13
7
6
11
10
15
8
3
12
Laboratory
En-Scores: S2 - Toluene
E_n-Score
4
3
2
1
0
-1
-2
-3
-4
-1.7
-0.6
-0.5
-0.4
-0.3
-0.4
0.0
0.2
0.4
0.3
0.9
2.0
2.3
2.8
4
14
1
2
9
13
7
6
11
10
15
8
3
12
Laboratory

Figure 8

Table 13

Sample Details

Sample No.	S2
Matrix	River Water
Analyte	Ethylbenzene
Unit	µg/L

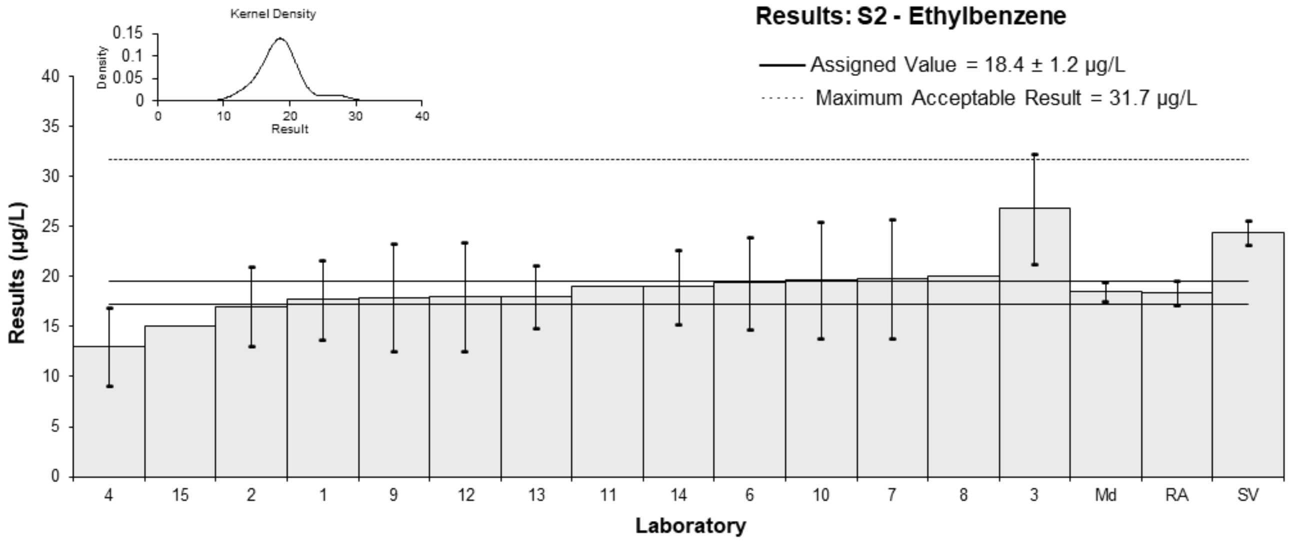
Participant Results

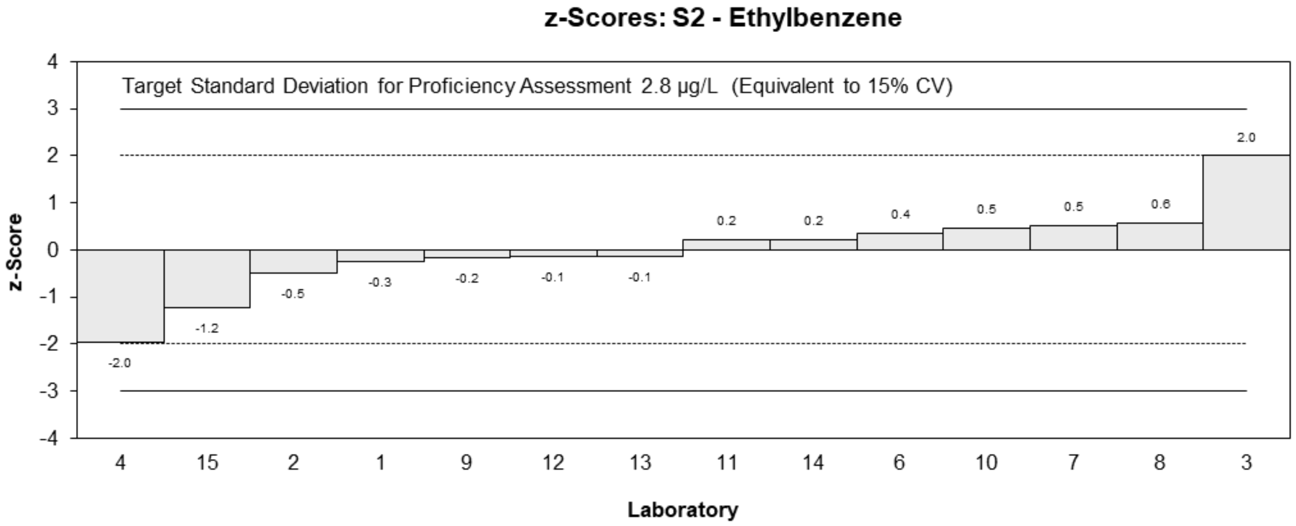
Lab. Code	Result	Uncertainty	z	E_n
1	17.7	4	-0.25	-0.17
2	17.04	3.92	-0.49	-0.33
3	26.8	5.5	2.00▼	
4	13	3.9	-1.96	-1.32
6	19.4	4.6	0.36	0.21
7	19.8	5.94	0.51	0.23
8	20	NR	0.58	1.33
9	17.9	5.37	-0.18	-0.09
10	19.7	5.8	0.47	0.22
11	19	NR	0.22	0.50
12	18	5.4	-0.14	-0.07
13	18	3.1	-0.14	-0.12
14	19	3.7	0.22	0.15
15	15	NR	-1.23	-2.83

▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	18.4	1.2
Spike Value	24.4	1.2
Robust Average	18.4	1.2
Max Acceptable Result	31.7	
Median	18.5	1.0
Mean	18.6	
N	14	
Max	26.8	
Min	13	
Robust SD	1.8	
Robust CV	9.6%	


Results: S2 - Ethylbenzene


z-Scores: S2 - Ethylbenzene

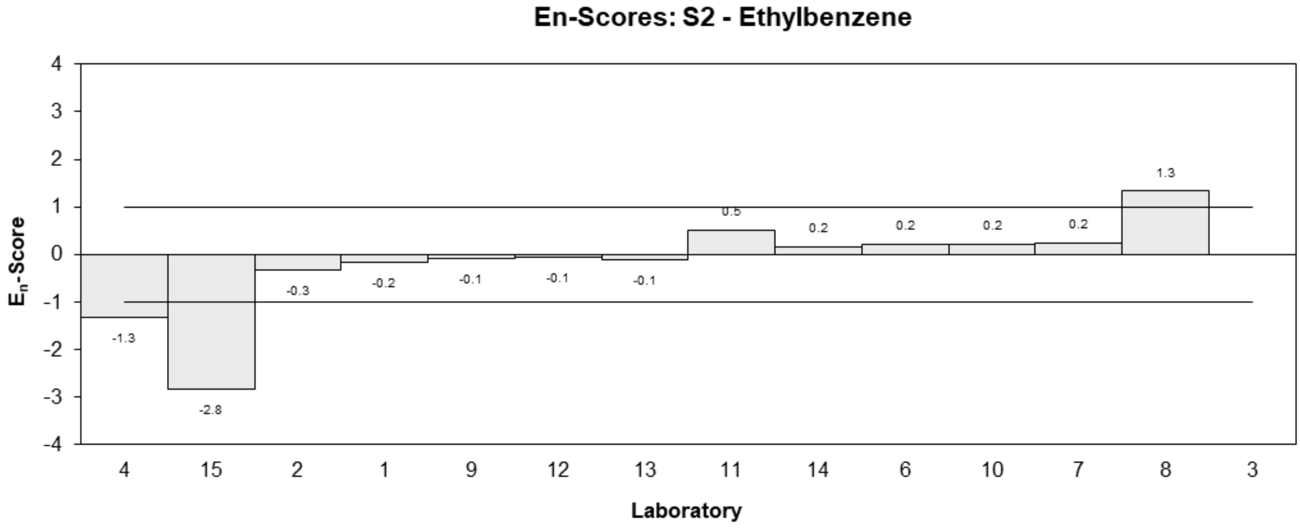

En-Scores: S2 - Ethylbenzene

Figure 9

Table 14

Sample Details

Sample No.	S2
Matrix	River Water
Analyte	Xylenes
Unit	µg/L

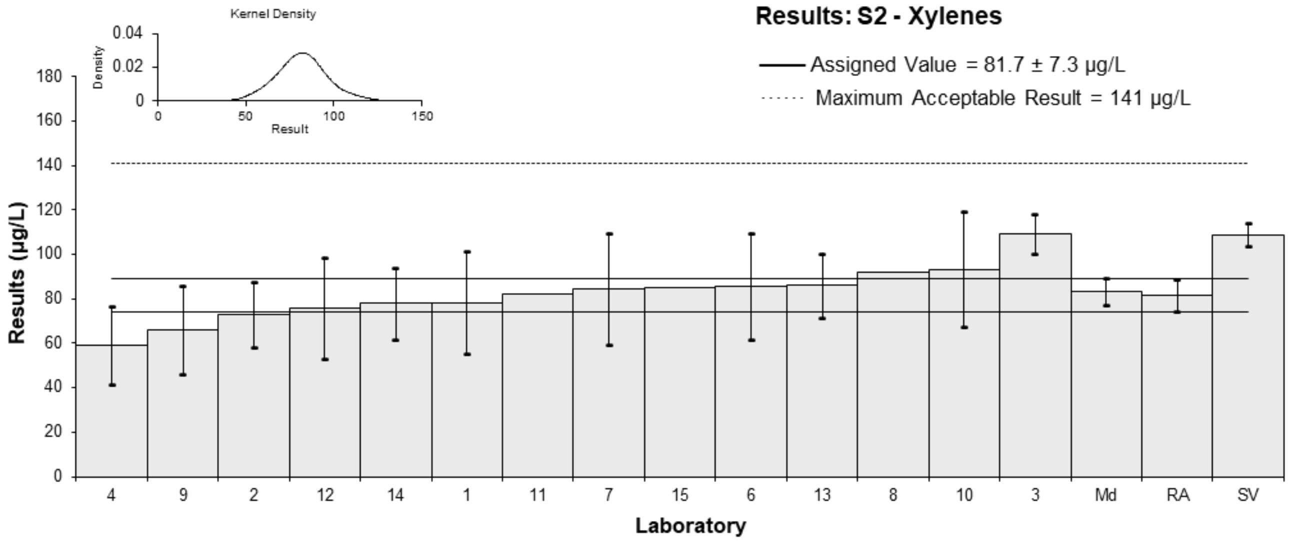
Participant Results

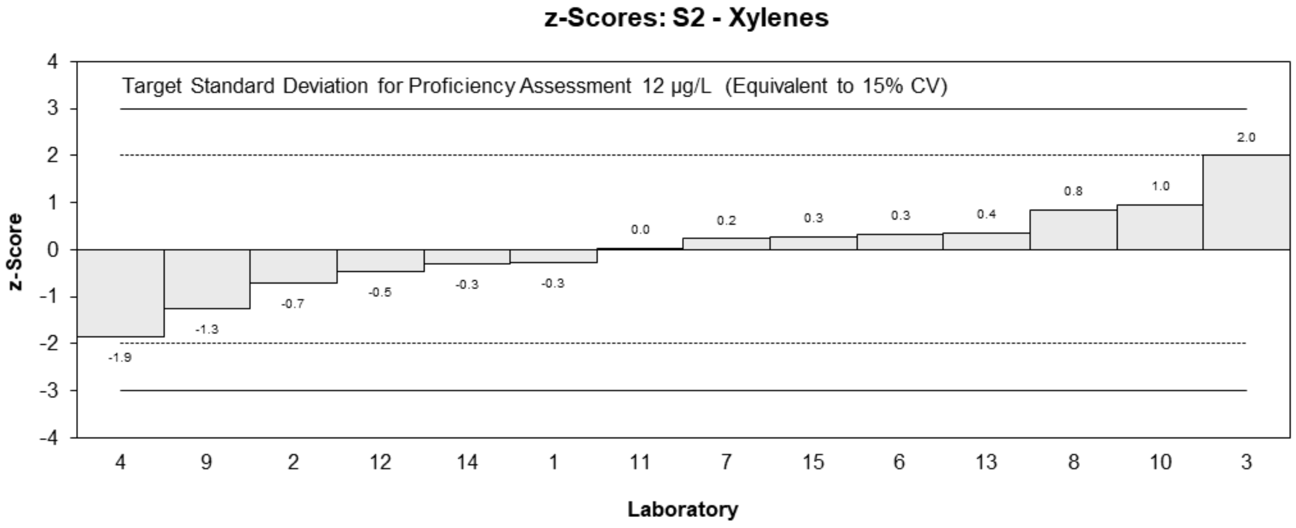
Lab. Code	Result	Uncertainty	z	E_n
1	78.3	23.1	-0.28	-0.14
2	72.9	14.58	-0.72	-0.54
3	109.3	8.8	2.00▼	
4	59	17.7	-1.85	-1.19
6	85.6	24.1	0.32	0.15
7	84.5	25.35	0.23	0.11
8	92	NR	0.84	1.41
9	66.2	19.86	-1.26	-0.73
10	93.4	26.1	0.95	0.43
11	82	NR	0.02	0.04
12	76	22.8	-0.47	-0.24
13	86	14.6	0.35	0.26
14	78	16	-0.30	-0.21
15	85	NR	0.27	0.45

▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	81.7	7.3
Spike Value	109	5
Robust Average	81.7	7.3
Max Acceptable Result	141	
Median	83.3	6.2
Mean	82.0	
N	14	
Max	109.3	
Min	59	
Robust SD	11	
Robust CV	13%	


Results: S2 - Xylenes
Assigned Value = 81.7 ± 7.3 µg/L
Maximum Acceptable Result = 141 µg/L
Kernel Density
Density
0.04
0.02
0
0
50
100
150
Result
Results (µg/L)
180
160
140
120
100
80
60
40
20
0
4
9
2
12
14
1
11
7
15
6
13
8
10
3
Md
RA
SV
Laboratory


z-Scores: S2 - Xylenes
Target Standard Deviation for Proficiency Assessment 12 µg/L (Equivalent to 15% CV)
z-Score
4
3
2
1
0
-1
-2
-3
-4
-1.9
-1.3
-0.7
-0.5
-0.3
-0.3
0.0
0.2
0.3
0.3
0.4
0.8
1.0
2.0
4
9
2
12
14
1
11
7
15
6
13
8
10
3
Laboratory

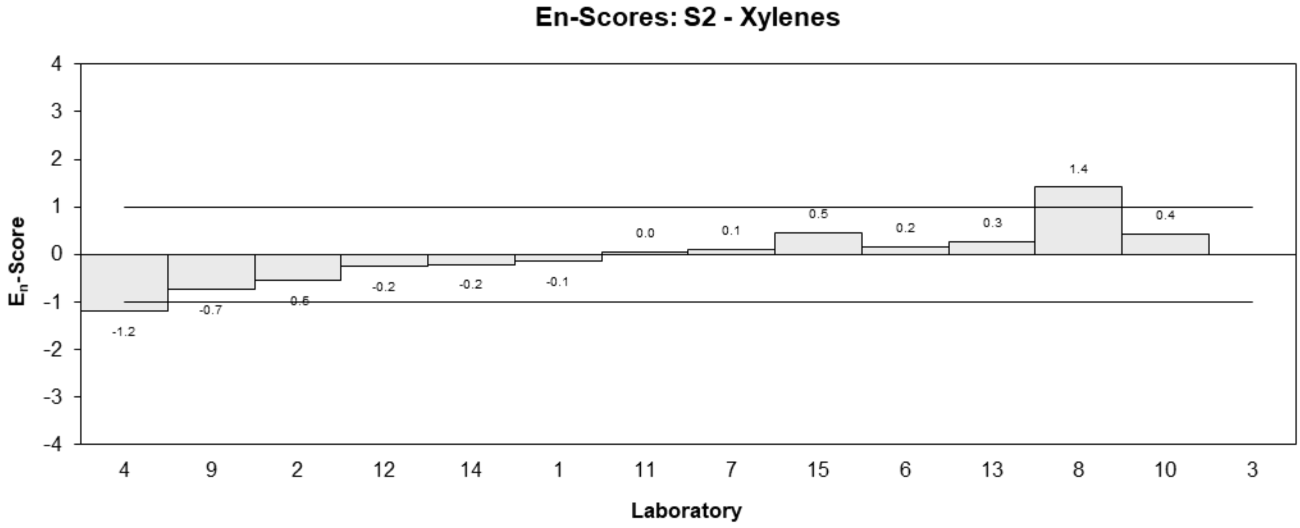

En-Scores: S2 - Xylenes
En-Score
4
3
2
1
0
-1
-2
-3
-4
-1.2
-0.7
-0.5
-0.2
-0.2
-0.1
0.0
0.1
0.5
0.2
0.3
1.4
0.4
4
9
2
12
14
1
11
7
15
6
13
8
10
3
Laboratory

Figure 10

Table 15

Sample Details

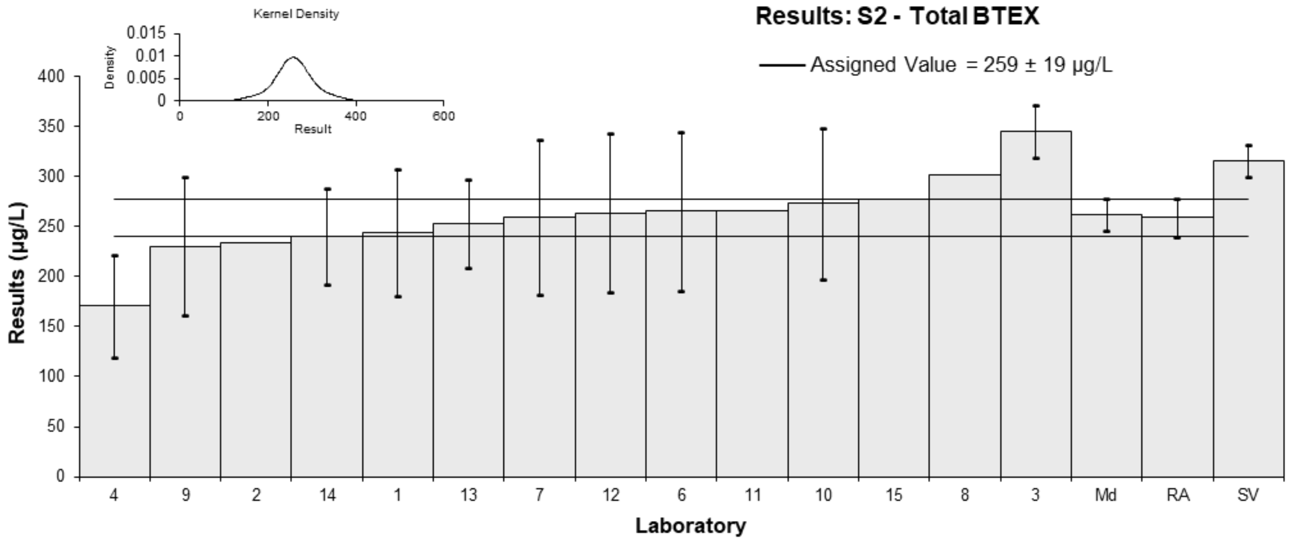
Sample No.	S2
Matrix	River Water
Analyte	Total BTEX
Unit	µg/L

Participant Results

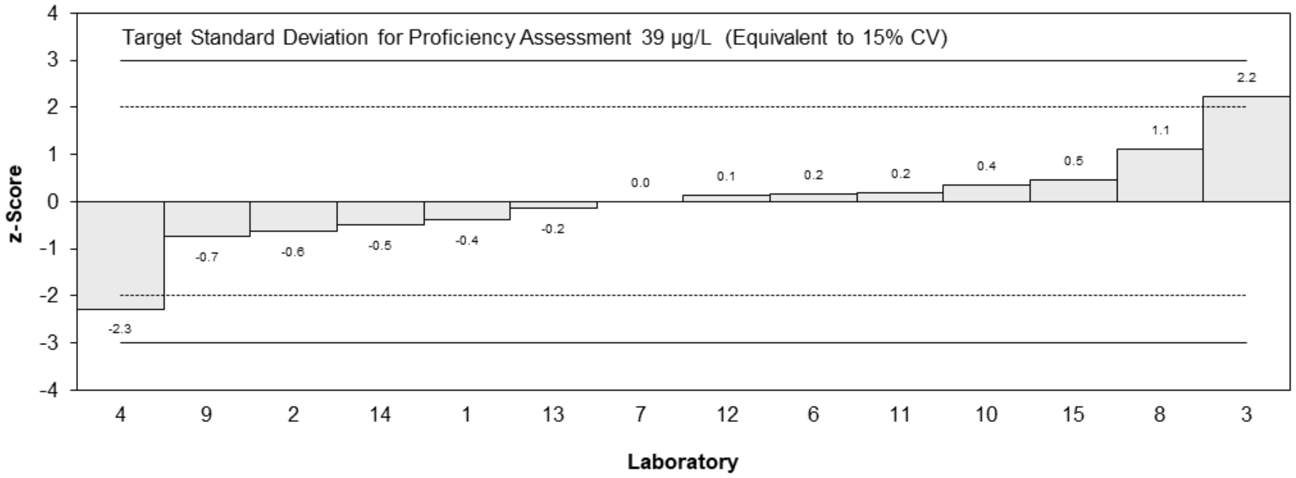
Lab. Code	Result	Uncertainty	z	E_n
1	244	63.8	-0.39	-0.23
2	234.2	NR	-0.64	-1.31
3	345.4	26.6	2.22	2.64
4	170.5	51.15	-2.28	-1.62
6	265.5	79.7	0.17	0.08
7	259.1	77.73	0.00	0.00
8	302	NR	1.11	2.26
9	230.4	69.12	-0.74	-0.40
10	273	75	0.36	0.18
11	266	NR	0.18	0.37
12	264	79.2	0.13	0.06
13	253	43.9	-0.15	-0.13
14	240	48	-0.49	-0.37
15	277	NR	0.46	0.95

Statistics

Assigned Value	259	19
Spike Value	316	16
Robust Average	259	19
Median	262	16
Mean	259	
N	14	
Max	345.4	
Min	170.5	
Robust SD	28	
Robust CV	11%	



z-Scores: S2 - Total BTEX



En-Scores: S2 - Total BTEX

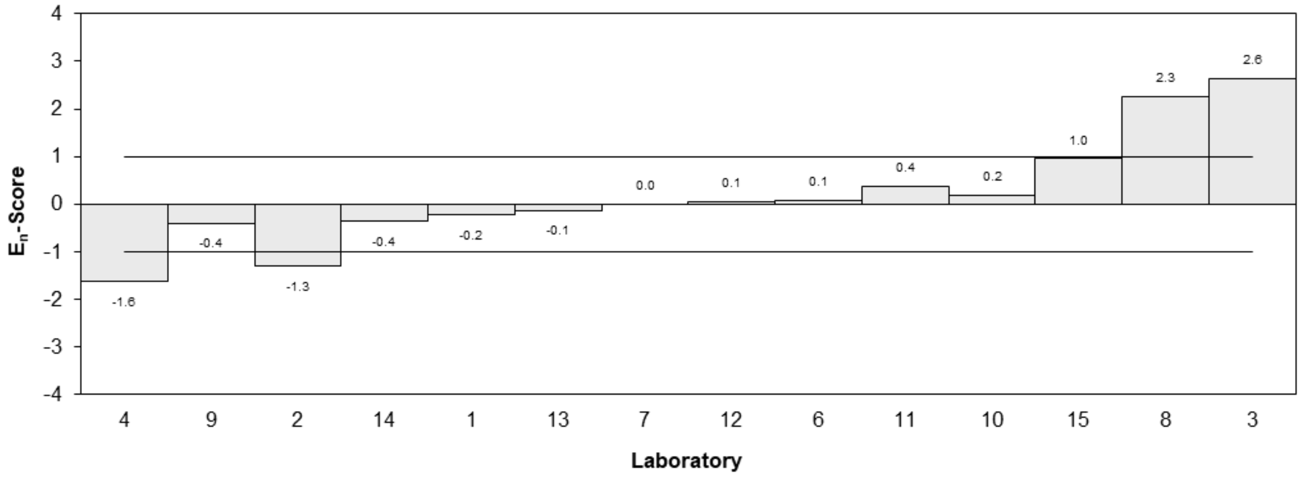


Figure 11

Table 16

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Acenaphthene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1**	3.8	1.3
2	7.97	1.99
3**	8.19	0.62
4	8.3	2.49
6	9.18	2.75
7	9.4	2.82
8	4.7	NR
9	9.6891	2.9067
10	8.15	2.61
11	4.55	NR
12	8.5	2.55
13	NS	NS
14	7.4	3.0
15	6.41	NR

** Excluded Result, see Section 4.2

Statistics

Assigned Value	Not Set	
Spike Value	15.9	0.8
Robust Average	7.7	1.4
Median	8.2	1.2
Mean	7.7	
N	11	
Max	9.6891	
Min	4.55	
Robust SD	1.9	
Robust CV	25%	

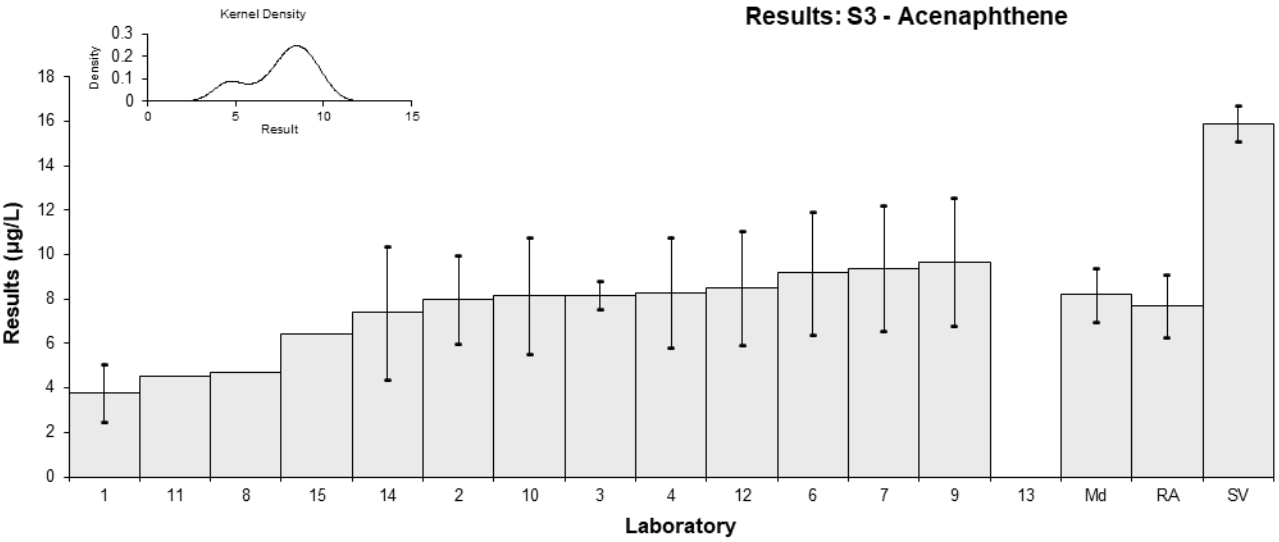

Results: S3 - Acenaphthene
Kernel Density
Density
0.3
0.2
0.1
0
0
5
10
15
Result
Results (µg/L)
18
16
14
12
10
8
6
4
2
0
1
11
8
15
14
2
10
3
4
12
6
7
9
13
Md
RA
SV
Laboratory

Figure 12

Table 17

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Anthracene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1**	1.8	0.7	-2.92	-1.83
2	3.27	0.82	0.15	0.08
3**	5.31	0.97	2.00▼	
4	3.3	0.99	0.21	0.10
6	2.64	0.79	-1.17	-0.66
7	2.6	0.78	-1.25	-0.71
8	3.18	NR	-0.04	-0.06
9	3.4665	1.04	0.56	0.25
10	2.85	0.8	-0.73	-0.41
11	3.42	NR	0.46	0.71
12	3.2	0.96	0.00	0.00
13	NS	NS		
14	3.6	1.4	0.83	0.28
15	3.63	NR	0.90	1.39

** Excluded Result, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	3.20	0.31
Spike Value	4.48	0.22
Robust Average	3.20	0.31
Max Acceptable Result	5.82	
Median	3.27	0.22
Mean	3.20	
N	11	
Max	3.63	
Min	2.6	
Robust SD	0.41	
Robust CV	13%	

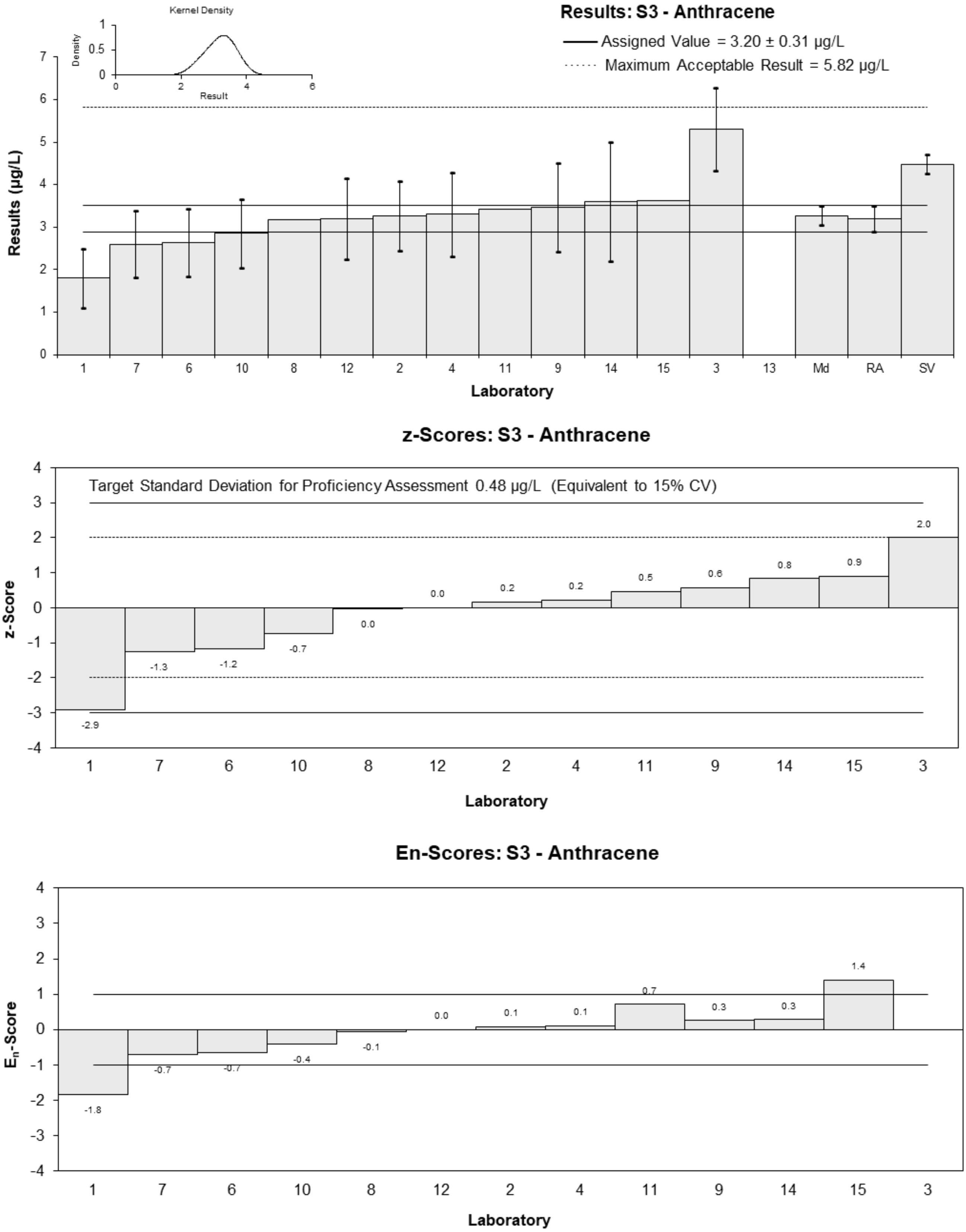

Results: S3 - Anthracene
Assigned Value = 3.20 ± 0.31 µg/L
Maximum Acceptable Result = 5.82 µg/L
Kernel Density
Density
Result
Results (µg/L)
Laboratory
1
7
6
10
8
12
2
4
11
9
14
15
3
13
Md
RA
SV
z-Scores: S3 - Anthracene
Target Standard Deviation for Proficiency Assessment 0.48 µg/L (Equivalent to 15% CV)
z-Score
-2.9
-1.3
-1.2
-0.7
0.0
0.0
0.2
0.2
0.5
0.6
0.8
0.9
2.0
En-Scores: S3 - Anthracene
E_n-Score
-1.8
-0.7
-0.7
-0.4
-0.1
0.0
0.1
0.1
0.7
0.3
0.3
1.4

Figure 13

Table 18

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Benz[*a*]anthracene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty
1**	1.9	0.6
2	3.47	0.87
3**	7.49	0.83
4	1.95	0.585
6	4.88	1.46
7	4.9	1.47
8	1.76	NR
9	3.3237	0.9971
10	5.65	1.75
11	1.18	NR
12	3.6	1.08
13	NS	NS
14	4.7	1.9
15	2.03	NR

** Excluded Result, see Section 4.2

Statistics

Assigned Value	Not Set	
Spike Value	4.50	0.22
Robust Average	3.4	1.3
Median	3.5	1.6
Mean	3.40	
N	11	
Max	5.65	
Min	1.18	
Robust SD	1.7	
Robust CV	50%	

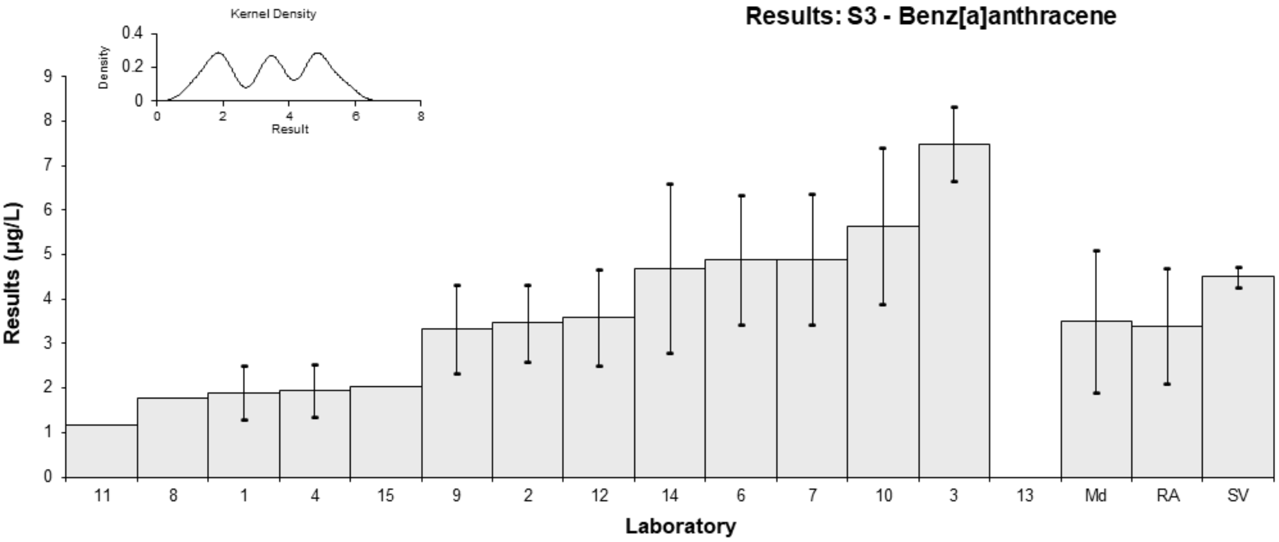

Results: S3 - Benz[a]anthracene
Kernel Density
Density
0.4
0.2
0
0
2
4
6
8
Result
Results (µg/L)
0
1
2
3
4
5
6
7
8
9
11
8
1
4
15
9
2
12
14
6
7
10
3
13
Md
RA
SV
Laboratory

Figure 14

Table 19

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Benzo[*a*]pyrene
Unit	µg/L

Participant Results

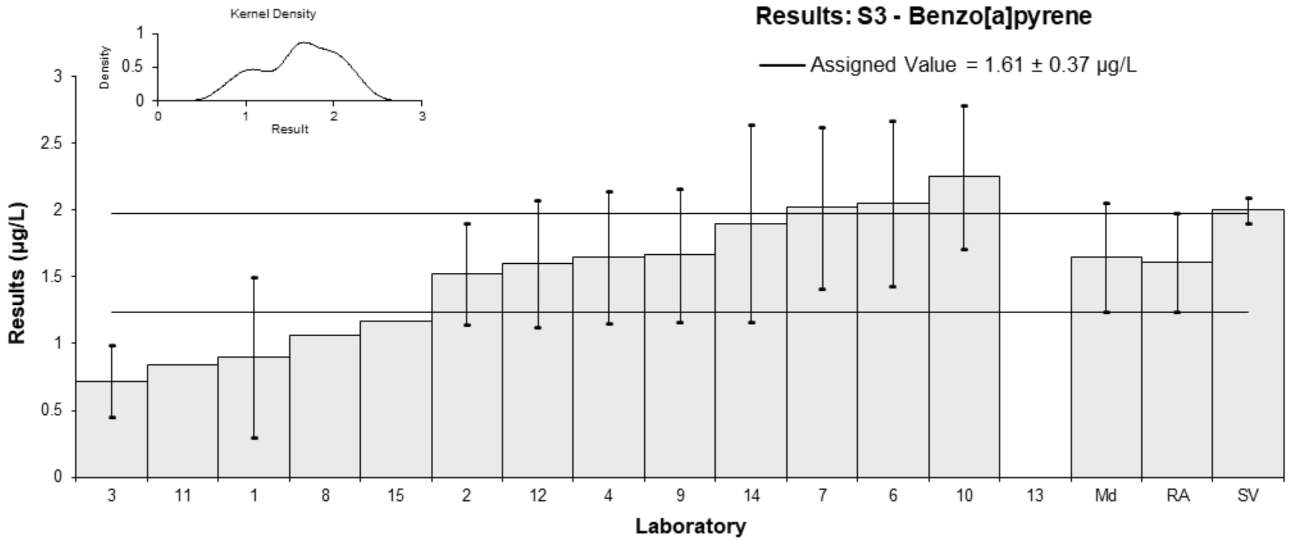
Lab. Code	Result	Uncertainty	z	E_n
1**	0.9	0.6	-2.94	-1.01
2	1.52	0.38	-0.37	-0.17
3**	0.72	0.27	-3.69	-1.94
4	1.65	0.495	0.17	0.06
6	2.05	0.62	1.82	0.61
7	2.02	0.606	1.70	0.58
8	1.06	NR	-2.28	-1.49
9	1.6648	0.4994	0.23	0.09
10	2.25	0.54	2.65	0.98
11	0.84	NR	-3.19	-2.08
12	1.6	0.48	-0.04	-0.02
13	NS	NS		
14	1.9	0.74	1.20	0.35
15	1.17	NR	-1.82	-1.19

** Excluded Result, see Section 4.2

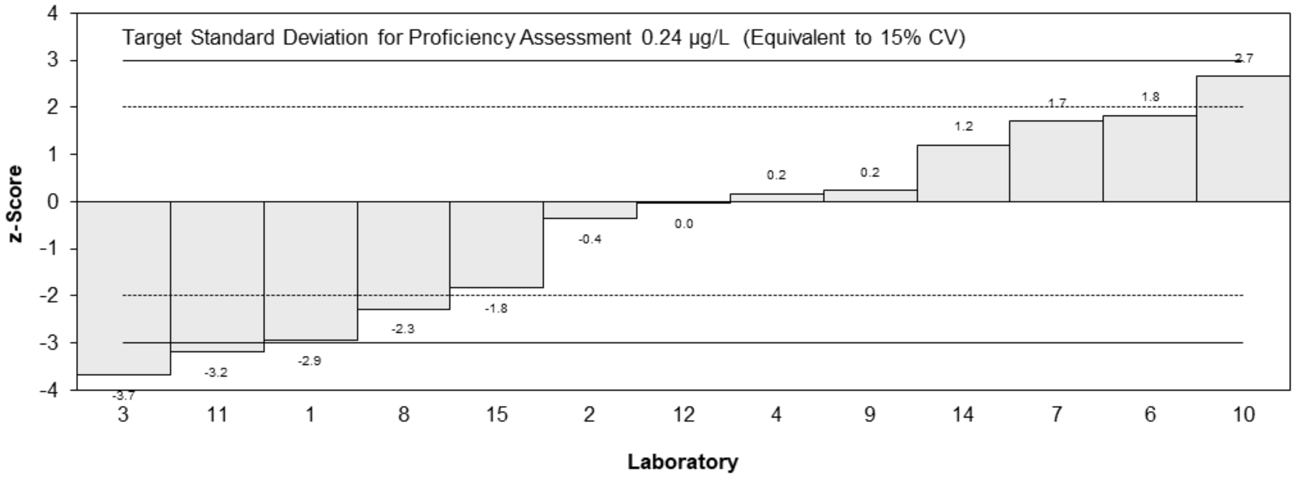
Statistics

Assigned Value	1.61	0.37
Spike Value	2.00	0.10
Robust Average	1.61	0.37
Median	1.65	0.41
Mean	1.61	
N	11	
Max	2.25	
Min	0.84	
Robust SD	0.50	
Robust CV	31%	

Results: S3 - Benzo[a]pyrene



z-Scores: S3 - Benzo[a]pyrene



En-Scores: S3 - Benzo[a]pyrene

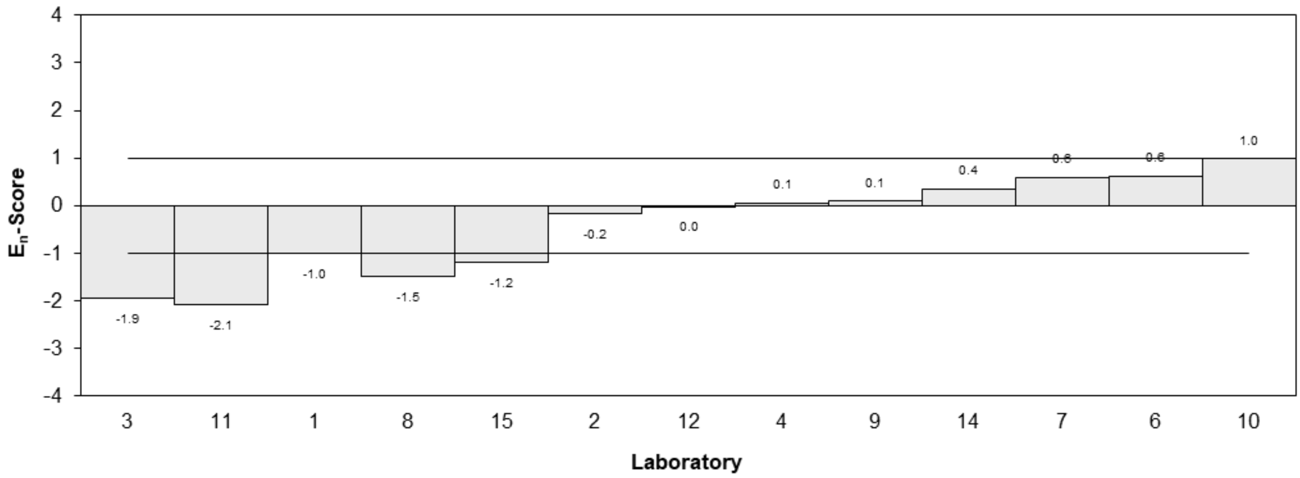


Figure 15

Table 20

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Chrysene
Unit	µg/L

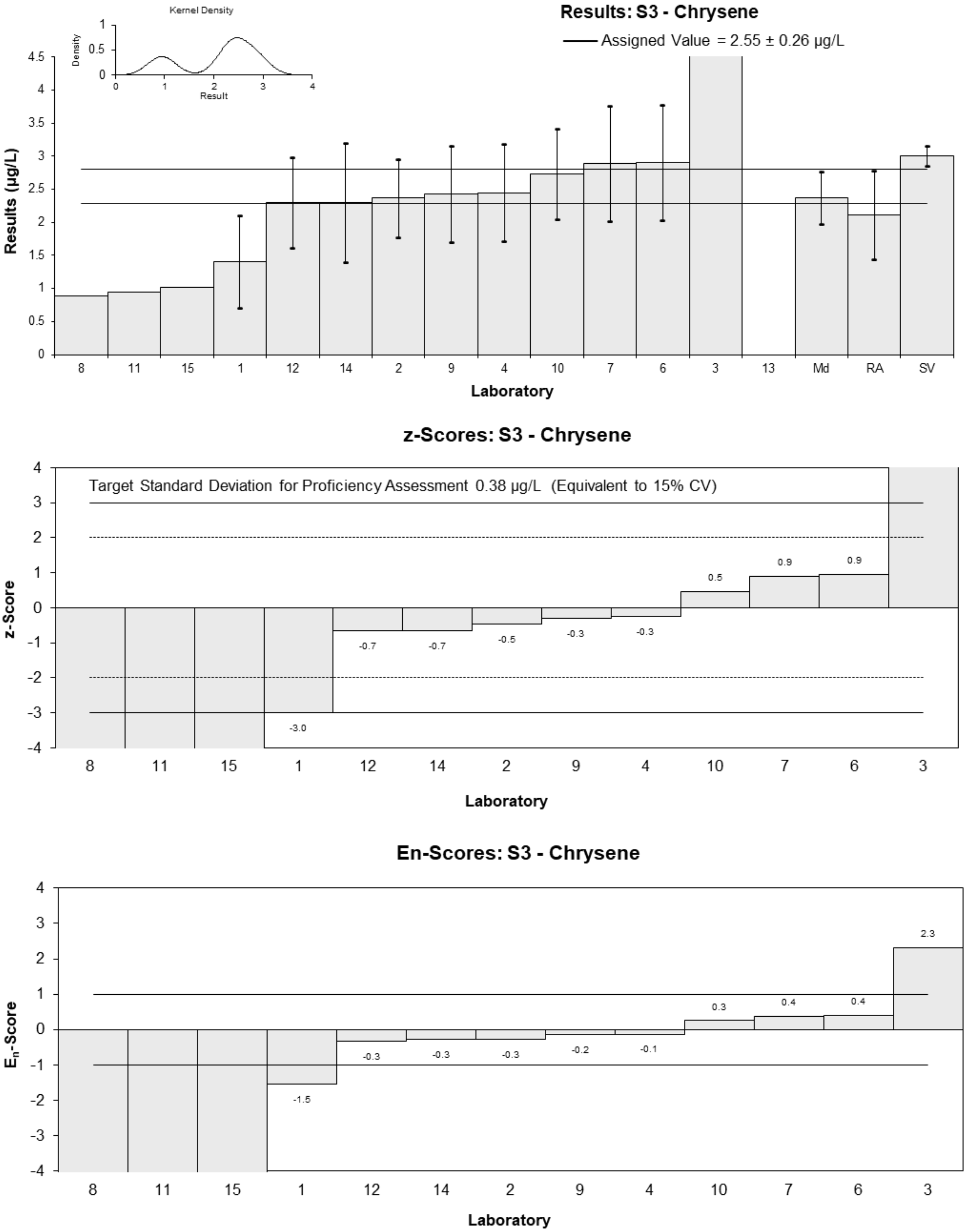
Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1**	1.4	0.7	-3.01	-1.54
2	2.37	0.59	-0.47	-0.28
3**	9.59	3.04	18.41	2.31
4	2.45	0.735	-0.26	-0.13
6	2.91	0.87	0.94	0.40
7	2.89	0.867	0.89	0.38
8*	0.88	NR	-4.37	-6.42
9	2.4319	0.7296	-0.31	-0.15
10	2.73	0.68	0.47	0.25
11*	0.94	NR	-4.21	-6.19
12	2.3	0.69	-0.65	-0.34
13	NS	NS		
14	2.3	0.9	-0.65	-0.27
15*	1.01	NR	-4.03	-5.92

* Outlier, ** Excluded Result, see Section 4.2

Statistics

Assigned Value	2.55	0.26
Spike Value	3.01	0.15
Robust Average	2.11	0.67
Median	2.37	0.40
Mean	2.11	
N	11	
Max	2.91	
Min	0.88	
Robust SD	0.88	
Robust CV	42%	



Results: S3 - Chrysene
Assigned Value = 2.55 ± 0.26 µg/L
Kernel Density
Density
Result
Results (µg/L)
Laboratory
8
11
15
1
12
14
2
9
4
10
7
6
3
13
Md
RA
SV
z-Scores: S3 - Chrysene
Target Standard Deviation for Proficiency Assessment 0.38 µg/L (Equivalent to 15% CV)
z-Score
-3.0
-0.7
-0.7
-0.5
-0.3
-0.3
0.5
0.9
0.9
Laboratory
En-Scores: S3 - Chrysene
E_n-Score
-1.5
-0.3
-0.3
-0.3
-0.2
-0.1
0.3
0.4
0.4
2.3
Laboratory

Figure 16

Table 21

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Fluoranthene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1**	1.4	0.6	-2.59	-1.32
2	2.16	0.54	-0.38	-0.21
3**	3.83	0.59	2.00▼	
4	2.4	0.72	0.32	0.14
6	2.33	0.70	0.12	0.05
7	2.42	0.726	0.38	0.16
8	1.62	NR	-1.95	-2.16
9	2.8009	0.8403	1.49	0.57
10	2.69	0.81	1.16	0.46
11	1.95	NR	-0.99	-1.10
12	2.5	0.75	0.61	0.26
13	NS	NS		
14	2.5	1.0	0.61	0.20
15	1.8	NR	-1.43	-1.58

** Excluded Result, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	2.29	0.31
Spike Value	3.00	0.15
Robust Average	2.29	0.31
Max Acceptable Result	3.89	
Median	2.40	0.27
Mean	2.29	
N	11	
Max	2.8009	
Min	1.62	
Robust SD	0.40	
Robust CV	18%	

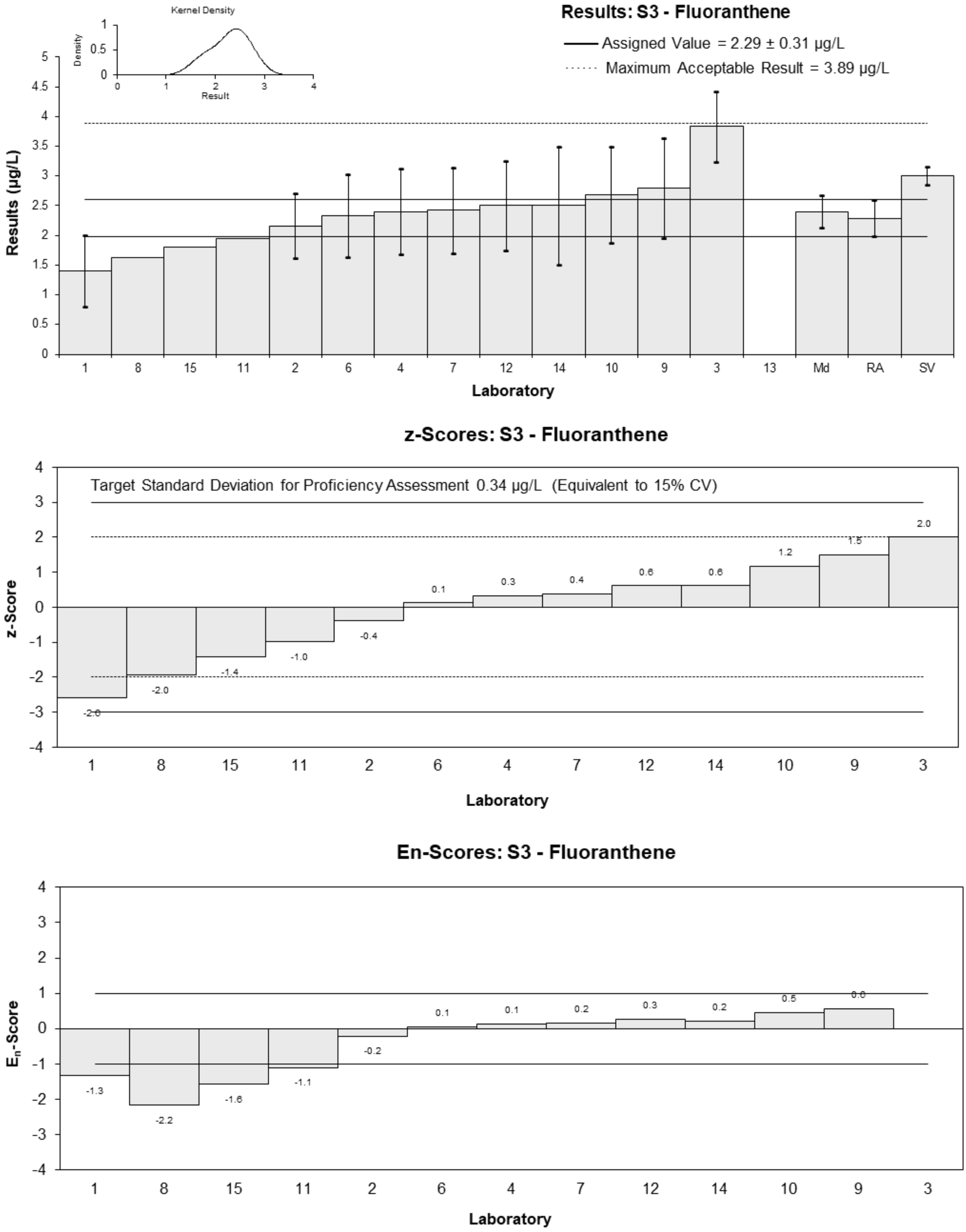

Kernel Density
Density
Result
Results: S3 - Fluoranthene
Assigned Value = 2.29 ± 0.31 µg/L
Maximum Acceptable Result = 3.89 µg/L
Results (µg/L)
Laboratory
1
8
15
11
2
6
4
7
12
14
10
9
3
13
Md
RA
SV
z-Scores: S3 - Fluoranthene
Target Standard Deviation for Proficiency Assessment 0.34 µg/L (Equivalent to 15% CV)
z-Score
-2.6
-2.0
-1.4
-1.0
-0.4
0.1
0.3
0.4
0.6
0.6
1.2
1.5
2.0
Laboratory
En-Scores: S3 - Fluoranthene
E_n-Score
-1.3
-2.2
-1.6
-1.1
-0.2
0.1
0.1
0.2
0.3
0.2
0.5
0.6
Laboratory

Figure 17

Table 22

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Fluorene
Unit	µg/L

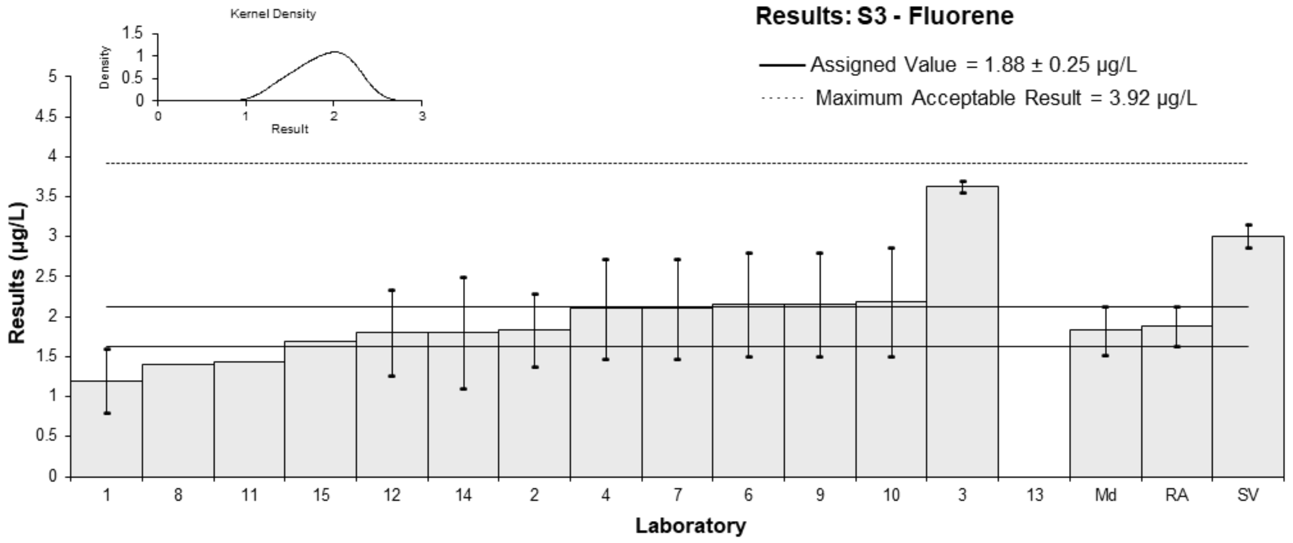
Participant Results

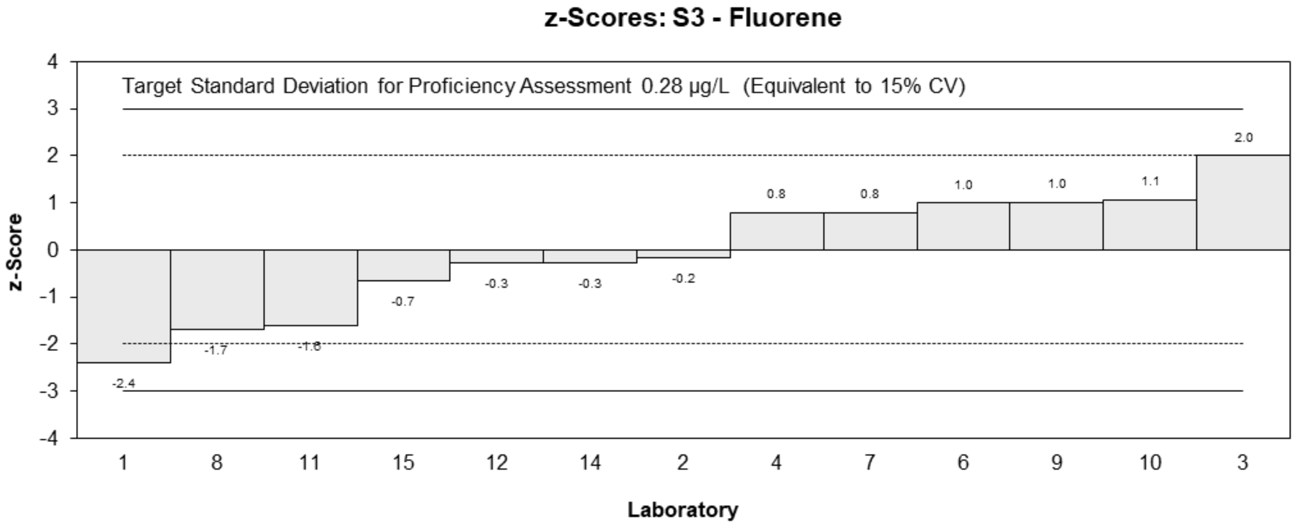
Lab. Code	Result	Uncertainty	z	E_n
1**	1.2	0.4	-2.41	-1.44
2	1.83	0.46	-0.18	-0.10
3**	3.63	0.07	2.00▼	
4	2.1	0.63	0.78	0.32
6	2.16	0.65	0.99	0.40
7	2.1	0.63	0.78	0.32
8	1.4	NR	-1.70	-1.92
9	2.1621	0.6486	1.00	0.41
10	2.18	0.68	1.06	0.41
11	1.43	NR	-1.60	-1.80
12	1.8	0.54	-0.28	-0.13
13	NS	NS		
14	1.8	0.7	-0.28	-0.11
15	1.69	NR	-0.67	-0.76

** Excluded Result, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	1.88	0.25
Spike Value	3.01	0.15
Robust Average	1.88	0.25
Max Acceptable Result	3.92	
Median	1.83	0.30
Mean	1.88	
N	11	
Max	2.18	
Min	1.4	
Robust SD	0.33	
Robust CV	17%	


Results: S3 - Fluorene
Assigned Value = 1.88 ± 0.25 µg/L
Maximum Acceptable Result = 3.92 µg/L
Kernel Density
Density
1.5
1
0.5
0
0
1
2
3
Result
Results (µg/L)
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1
8
11
15
12
14
2
4
7
6
9
10
3
13
Md
RA
SV
Laboratory


z-Scores: S3 - Fluorene
Target Standard Deviation for Proficiency Assessment 0.28 µg/L (Equivalent to 15% CV)
z-Score
4
3
2
1
0
-1
-2
-3
-4
-2.4
-1.7
-1.6
-0.7
-0.3
-0.3
-0.2
0.8
0.8
1.0
1.0
1.1
2.0
1
8
11
15
12
14
2
4
7
6
9
10
3
Laboratory

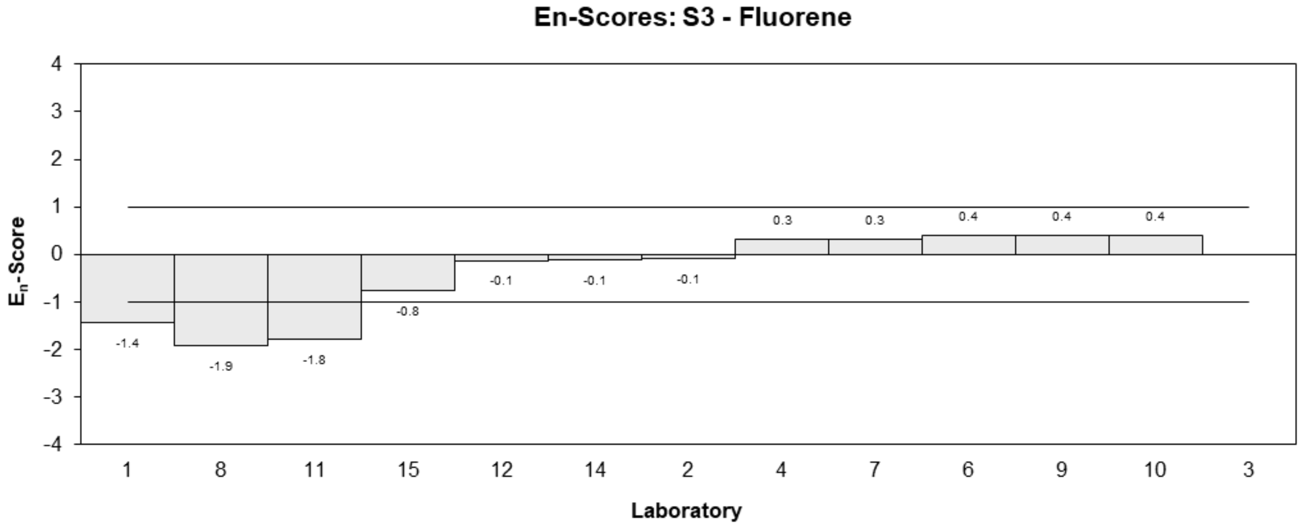

En-Scores: S3 - Fluorene
E_n-Score
4
3
2
1
0
-1
-2
-3
-4
-1.4
-1.9
-1.8
-0.8
-0.1
-0.1
-0.1
0.3
0.3
0.4
0.4
0.4
1
8
11
15
12
14
2
4
7
6
9
10
3
Laboratory

Figure 18

Table 23

Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Pyrene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1**	0.9	0.5	-2.92	-1.19
2	1.37	0.34	-0.96	-0.50
3**	4.83	0.18	13.46	9.01
4	1.7	0.51	0.42	0.17
6	2.05	0.62	1.87	0.65
7	2.02	0.606	1.75	0.62
8	1.01	NR	-2.46	-1.90
9	1.8062	0.5419	0.86	0.33
10	1.93	0.56	1.37	0.52
11	1.21	NR	-1.63	-1.26
12	1.7	0.51	0.42	0.17
13	NS	NS		
14	1.7	0.7	0.42	0.13
15	1.14	NR	-1.92	-1.48

** Excluded Result, see Section 4.2

Statistics

Assigned Value	1.60	0.31
Spike Value	2.00	0.10
Robust Average	1.60	0.31
Median	1.70	0.36
Mean	1.60	
N	11	
Max	2.05	
Min	1.01	
Robust SD	0.41	
Robust CV	26%	

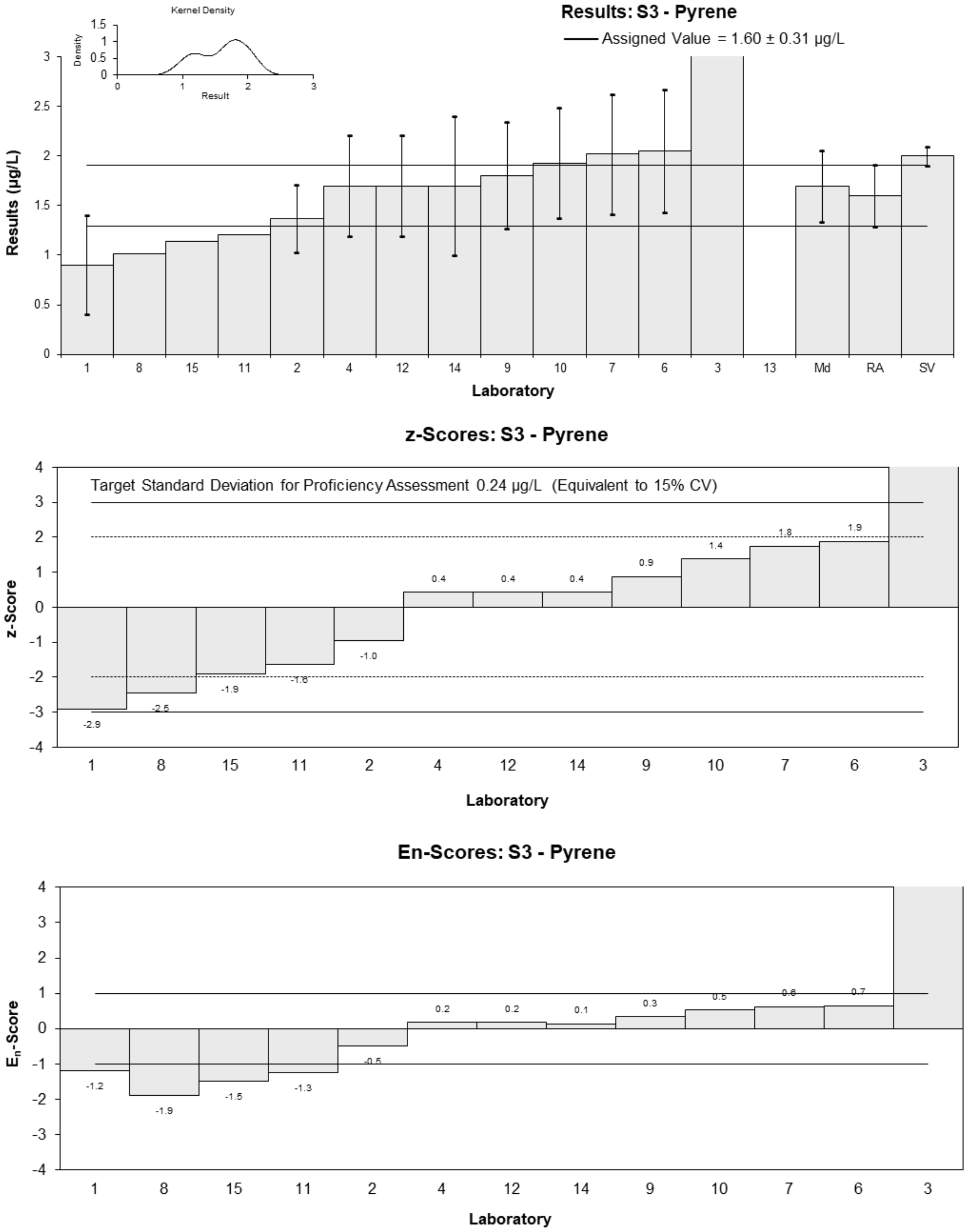

Kernel Density
Density
Result
Results: S3 - Pyrene
Assigned Value = 1.60 ± 0.31 μg/L
Results (μg/L)
Laboratory
1 8 15 11 2 4 12 14 9 10 7 6 3 13 Md RA SV
z-Scores: S3 - Pyrene
Target Standard Deviation for Proficiency Assessment 0.24 μg/L (Equivalent to 15% CV)
z-Score
-2.9 -2.5 -1.9 -1.6 -1.0 0.4 0.4 0.4 0.9 1.4 1.8 1.9
Laboratory
1 8 15 11 2 4 12 14 9 10 7 6 3
En-Scores: S3 - Pyrene
E_n-Score
-1.2 -1.9 -1.5 -1.3 -0.5 0.2 0.2 0.1 0.3 0.5 0.6 0.7
Laboratory
1 8 15 11 2 4 12 14 9 10 7 6 3

Figure 19

6 DISCUSSION OF RESULTS

6.1 Assigned Value

The assigned values for all scored analytes were the robust averages of participants' results. If there were results less than 50% or greater than 150% of the robust average, these were excluded from the calculation of each assigned value.[3,4] The robust averages and associated expanded uncertainties were calculated using the procedure described in ISO 13528.[7] The calculation of the expanded uncertainty for robust averages is presented in Appendix 3, using Sample S2 ethylbenzene as an example.

Traceability: The consensus of participants' results is not traceable to any external reference, so although expressed in SI units, metrological traceability has not been established.

No assigned value was set for analytes in Sample S1, or for benz[*a*]anthracene in Sample S3, as there was no consensus between results reported by participants. No assigned value was set for Sample S3 acenaphthene as the consensus value was significantly lower than the spiked value, though there was a good consensus between participants' results. Sample S2 C6-C10 range was also not scored; historically this has been due to its volatile nature and therefore data has been provided for information only, though participants' results were in good agreement with each other for this study. For these analytes which were not scored, participants may still compare their results with the descriptive statistics and spiked values presented in Section 5.

A comparison of the assigned values and spiked values is presented in Table 24. Assigned values were set where the consensus value to spiked value ratio was similar to those observed in previous NMIA PT studies, and if there was a reasonable consensus of participants' results.

Table 24 Comparison of Assigned Value and Spiked Value

Sample	Analyte	Assigned Value (µg/L)	Spiked Value (µg/L)	Assigned Value / Spiked Value (%)
S1**	TRH	2200*	1900	116
S2	Benzene	53.7	55.1	97
	Toluene	105	128	82
	Ethylbenzene	18.4	24.4	75
	Xylenes	81.7	109	75
	Total BTEX	259	316	82
S3	Acenaphthene	7.7*	15.9	48
	Anthracene	3.20	4.48	71
	Benz[*a*]anthracene	3.4*	4.50	76
	Benzo[*a*]pyrene	1.61	2.00	81
	Chrysene	2.55	3.01	85
	Fluoranthene	2.29	3.00	76
	Fluorene	1.88	3.01	62
	Pyrene	1.60	2.00	80

* Robust average (assigned value not set).

** Sample S1 was spiked with diesel. Spiked value is approximate.

6.2 Measurement Uncertainty Reported by Participants

Participants were asked to report the expanded uncertainty evaluations associated with their results and the basis of this uncertainty. It is a requirement of ISO/IEC 17025 that laboratories have procedures to evaluate the uncertainty of chemical measurements and to report this uncertainty in specific circumstances, including when the client's instruction so requires.[9]

Of 222 numeric results submitted for analytes of interest in this study, 167 (75%) were reported with an associated uncertainty. Participants used a wide variety of procedures to evaluate their expanded MU (Table 3).

Laboratories **8**, **11** and **15** did not report any uncertainties. These participants noted that they were currently updating their uncertainty evaluation procedures. These participants reported being accredited to ISO/IEC 17025 for PAHs analysis only.

The Sample S1 TRH result from Laboratory **2** had no uncertainty as the result was calculated by the study coordinator by summing the individual hydrocarbon range results reported. This participant did not report an uncertainty for their Sample S2 Total BTEX result. This participant reported being accredited to ISO/IEC 17025 across all analyte types.

The magnitude of reported uncertainties was within the range of 1.9% to 67% relative. In general, an expanded measurement uncertainty of less than 10% relative is likely to be unrealistically small for the routine measurement of hydrocarbons in river water, while an expanded uncertainty of over 50% is likely too large to be fit for purpose. Of 167 MUs, 157 (94%) were between 10% and 50% relative, while eight were less than 10% relative, and two were greater than 50% relative.

Participants were requested to report the coverage factor associated with their uncertainty. All participants reporting a coverage factor reported using k = 2.

Uncertainties associated with results returning an acceptable *z*-score, but an unacceptable E_n-score, may have been underestimated.

An evaluation of uncertainty expressed as a value should not be attached to a non-value result.[10] Laboratories **7** and **10** attached an uncertainty to at least one of their non-value results.

In some cases, results were reported with an inappropriate number of significant figures. Including too many significant figures may inaccurately reflect the measurement precision. The recommended format is to write uncertainty to no more than two significant figures and then to write the result with the corresponding number of decimal places. For example, instead of 1.6648 ± 0.4994 µg/L, it is better to report 1.66 ± 0.50 µg/L.[10]

6.3 z-Score

Target SDs equivalent to 15% PCV were used to calculate *z*-scores for all scored analytes. CVs predicted by the Thompson-Horwitz equation,[8] between-laboratory CVs and target SDs (as PCV) for this study are presented for comparison in Table 25.

Table 25 Comparison of Thompson-Horwitz CV, Between-Laboratory CV and Target SD

Sample	Analyte	Assigned Value (µg/L)	Thompson-Horwitz CV* (%)	Between-Laboratory CV* (%)	Target SD (as PCV) (%)
S1	>C10-C16	730**	17	105	Not Set
	>C16-C34	420**	18	43	Not Set
	>C34-C40****	1000***	15	147	Not Set

Sample	Analyte	Assigned Value (µg/L)	Thompson-Horwitz CV* (%)	Between-Laboratory CV* (%)	Target SD (as PCV) (%)
	TRH	2200**	14	117	Not Set
S2	C6-C10	500**	18	46	Not Set
	Benzene	53.7	22	13	15
	Toluene	105	22	16	15
	Ethylbenzene	18.4	22	9.6	15
	Xylenes	81.7	22	13	15
	Total BTEX	259	20	11	15
S3	Acenaphthene	7.7**	22	25	Not Set
	Anthracene	3.20	22	13	15
	Benz[*a*]anthracene	3.4**	22	50	Not Set
	Benzo[*a*]pyrene	1.61	22	31	15
	Chrysene	2.55	22	11	15
	Fluoranthene	2.29	22	18	15
	Fluorene	1.88	22	17	15
	Pyrene	1.60	22	26	15

* Thompson-Horwitz CV calculated from the assigned value. The between-laboratory CV is the robust between-laboratory CV with outliers removed, if applicable.

** Robust average (assigned value not set).

*** Median (assigned value not set).

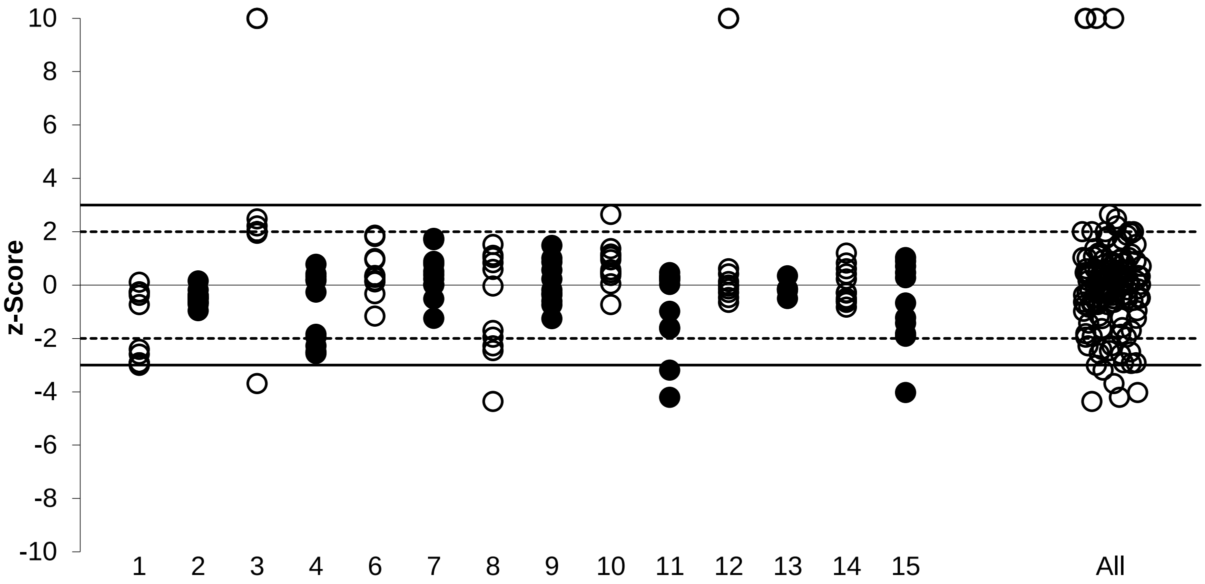
**** Analysis of the diesel used for spiking Sample S1 indicated no >C34-C40 component.

To account for possible low bias in the consensus values due to laboratories using inefficient extraction or analytical techniques, a total of five *z*-scores were adjusted across the following analytes: Sample S2 ethylbenzene and xylenes, Sample S3 anthracene, fluoranthene and fluorene. For these analytes, a maximum acceptable result was set as the spiked value plus two target SDs of the spiked value. Results lower than the maximum acceptable result but with a *z*-score greater than 2.0 had their *z*-score adjusted to 2.0. This ensured that participants reporting results close to the spiked value were not penalised. *z*-Scores for results greater than the maximum acceptable result, and *z*-scores less than 2.0, were left unaltered.

Of 148 results for which *z*-scores were calculated, 125 (84%) returned an acceptable score of $|z| \leq 2.0$.

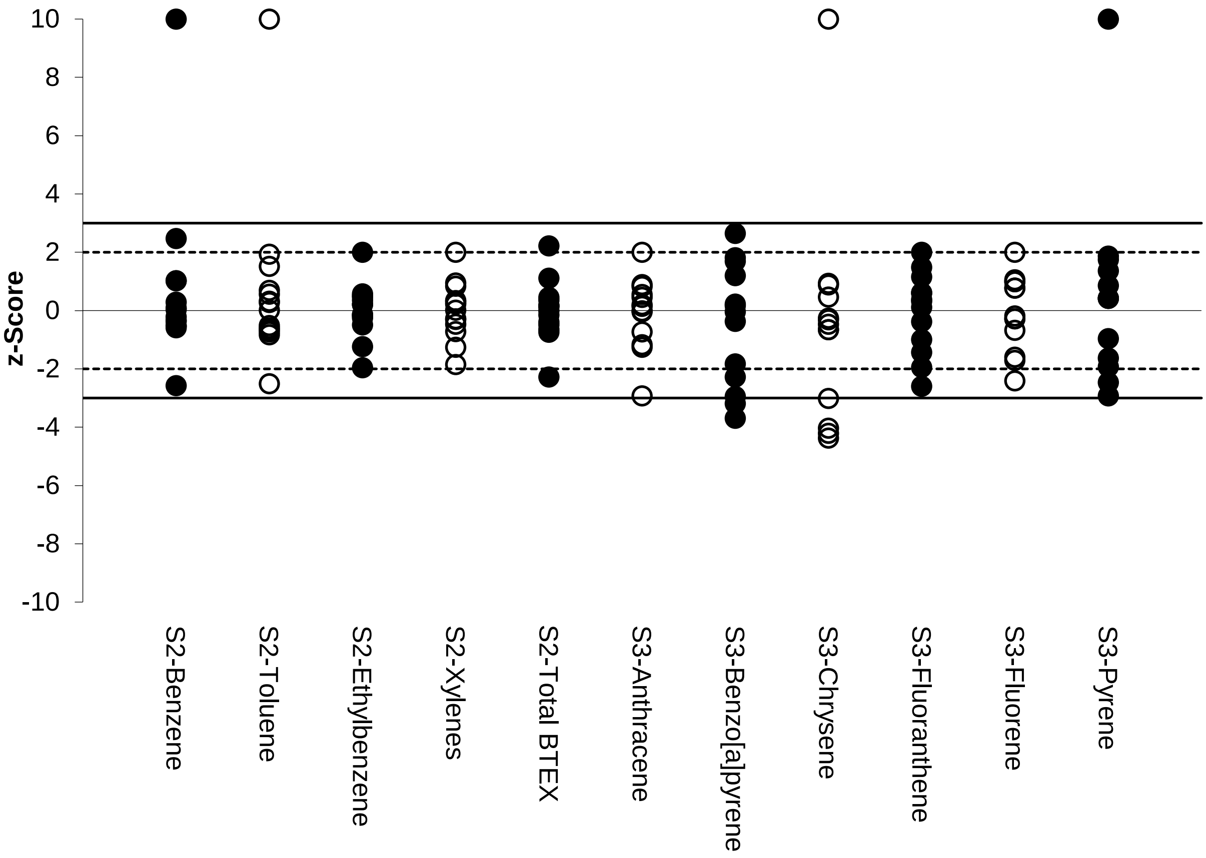
Laboratories **1**, **2**, **3**, **4**, **6**, **7**, **8**, **9**, **10**, **11**, **12**, **14** and **15** reported results for all 11 scored analytes. Of these participants, Laboratories **2**, **6**, **7**, **9** and **14** returned acceptable *z*-scores for all scored analytes. One other participant received acceptable *z*-scores for all their reported results that were scored: Laboratory **13** (5).

A summary of participants’ *z*-scores dispersal is presented by laboratory in Figure 20 and by analyte in Figure 21.



z-Scores greater than 10.0 have been plotted at 10.0.

Figure 20 *z*-Score Dispersal by Laboratory

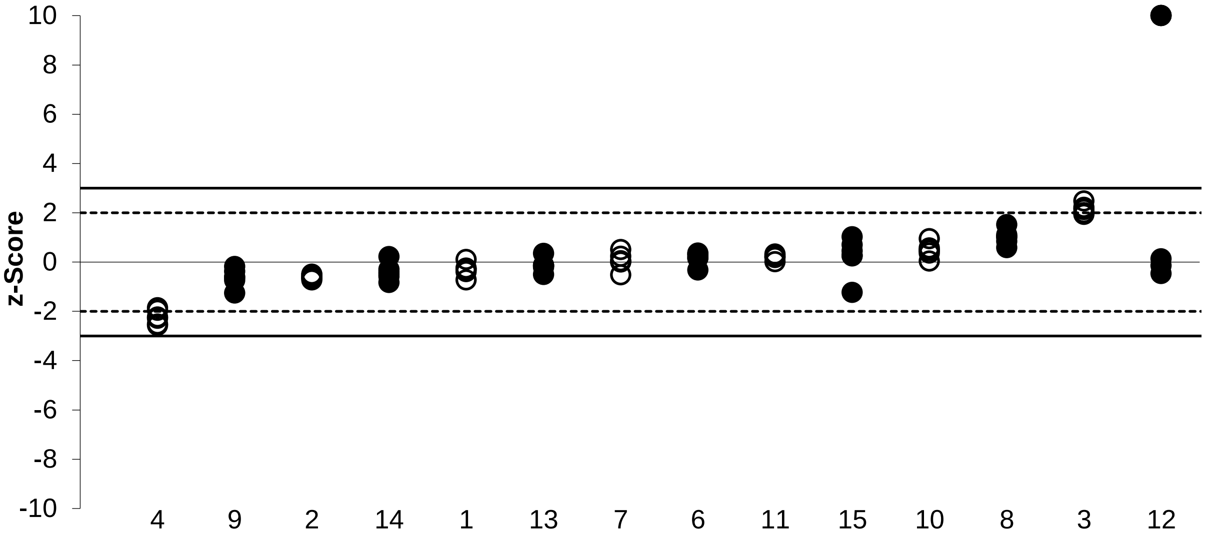


z-Scores greater than 10.0 have been plotted at 10.0.

Figure 21 *z*-Score Dispersal by Sample and Analyte

Figure 22 presents participants' *z*-scores for Sample S2 (BTEX) only. Participants with a trend of *z*-scores below the zero line possibly had an inefficient extraction process for BTEX. As the ratio of the assigned value to the spiked value was 82% for total BTEX, participants reporting results with higher acceptable *z*-scores may have more efficient extraction methodologies.

While most participants' results for BTEX were consistent with regards to bias, the results from Laboratory **12** were relatively varied. This participant returned three results with acceptable *z*-scores, and two results which were significantly higher than the assigned value (1423% and 603%).

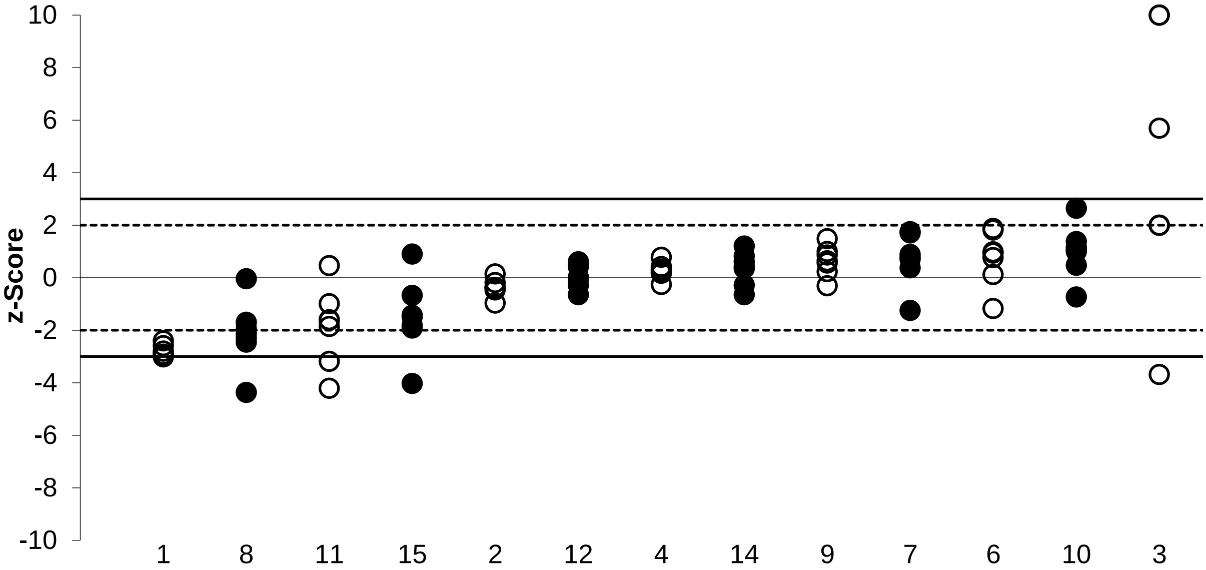


z-Scores greater than 10.0 have been plotted at 10.0.

Figure 22 Sample S2 (BTEX) *z*-Score Dispersal by Laboratory

Figure 23 presents participants' *z*-scores for Samples S3 (PAHs) only. Participants with a trend of *z*-scores below the zero line may have an inefficient extraction process for PAHs. As the ratio of the assigned values to the spiked values ranged from 62% to 85%, results with higher acceptable *z*-scores may correspond to the more efficient extraction of PAHs.

PAHs results from Laboratory **3** were extremely varied, with results ranging from 45% to 376% of the assigned value. With consideration also to the additional analytes reported by this participant (Section 6.6), this participant may have reported results for a different sample.



z-Scores greater than 10.0 have been plotted at 10.0.

Figure 23 Sample S3 (PAHs) *z*-Score Dispersal by Laboratory

6.4 E_n-Score

E_n-Scores can be interpreted in conjunction with *z*-scores, as an unacceptable E_n-score can either be caused by issues with measurement, or uncertainty, or both. If a participant did not report any uncertainty with a result, an expanded uncertainty of zero (0) was used to calculate the E_n-score. For results whose *z*-scores were adjusted as discussed in Section 6.3, no E_n-score has been calculated.

Of 143 results for which E_n-scores were calculated, 101 (71%) returned an acceptable score of $|E_n| < 1.0$, indicating agreement of the participant's result with the assigned value within their respective expanded uncertainties.

Laboratories **6**, **7**, **9**, **10** and **14** returned acceptable E_n-scores for all 11 scored analytes. One other participant received acceptable E_n-scores for all their reported results that were scored: Laboratory **13** (5).

Laboratory **3** did not return any acceptable E_n-scores.

A summary of E_n-score dispersal by laboratory is presented in Figure 24.

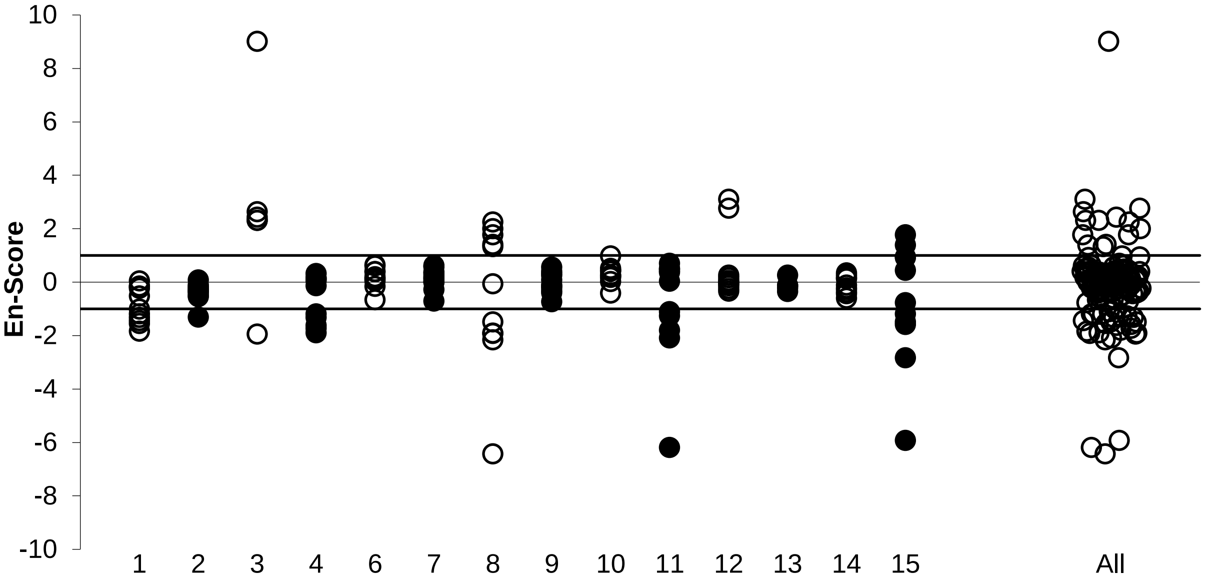


Figure 24 E_n-Score Dispersal by Laboratory

6.5 False Negatives

Table 26 presents false negative results. These are analytes present in the samples which a participant tested for but did not report a numeric result; for example, when participants reported a 'less than' result (< *x*) when the assigned value was greater than their limit of reporting (LOR), or if no value was reported. For analytes where no assigned value was set, results have only been considered to be false negatives where the consensus value and spiked value were significantly higher than the participants' LOR (i.e. the consensus value minus the expanded uncertainty, and the spiked value minus the expanded uncertainty, were both greater than the LOR), or if no value was reported.

Table 26 False Negatives

Lab. Code	Sample	Analyte	Assigned Value (µg/L)	Spiked Value (µg/L)	Result (µg/L)
14	S1	Total TRH	2200*	1900	<250

* Robust average (assigned value not set).

6.6 Reporting of Additional Analytes

Table 27 presents analytes reported by participants that were not spiked into the test samples by the study coordinator. In general, participants should take care to avoid any potential cross-contamination when analysing their samples.

Table 27 Non-Spiked Analytes Reported by Participants

Lab. Code	Sample	Analyte	Result (µg/L)	Uncertainty (µg/L)
3	S3	Naphthalene	2.33	0.67
		Acenaphthylene	2	0.38
		Phenanthrene	3.25	0.21
		Benzo[b]fluoranthene	4.33	1.07
		Indeno[1,2,3-cd]pyrene	0.68	0.46
		Dibenz[a,h]anthracene	0.91	0.34
		Benzo[g,h,i]perylene	1.06	0.32
8	S3	Naphthalene	0.32	NR
15	S3	Naphthalene	0.29	NR

6.7 Participants’ Analytical Methods

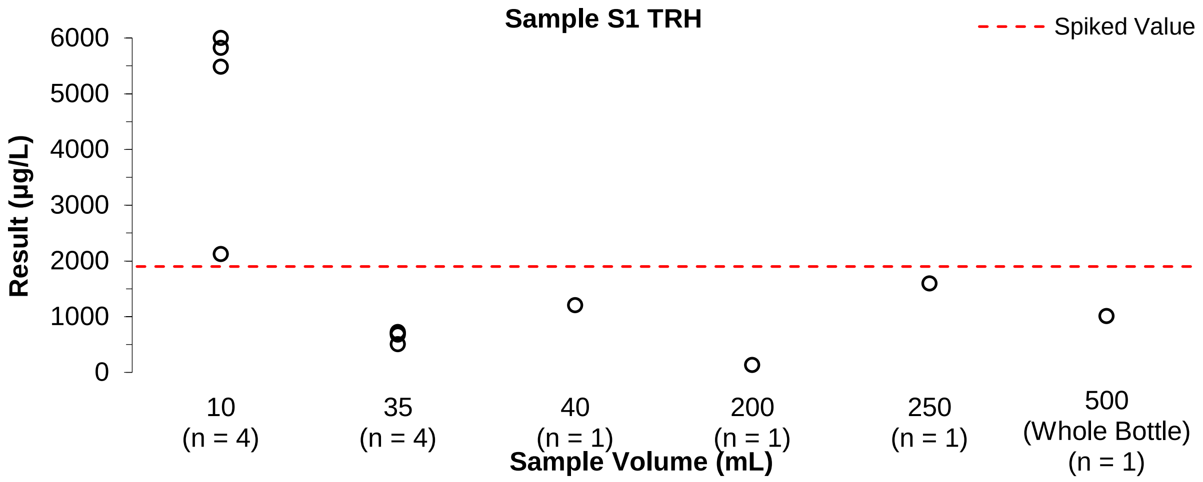
Participants’ results excluded from all summary statistics in Section 5 have also been excluded from discussion in this section. Where charts refer to ‘n = *x*’, this corresponds to *x* number of participants using that technique.

Sample S1 TRH

No assigned value was able to be set for Sample S1 TRH as there was no consensus between results reported by participants. Participant results have been compared to the spiked value in this section.

Of participants reporting numeric results, one participant reported taking the whole sample for analysis, while others reported sample test portions ranging from 10 mL to 250 mL.

A comparison of the TRH results and sample volume used is presented in Figure 25. Three participants using 10 mL for sample analysis reported extremely high values for TRH. Caution should be exercised when a small sample size is taken for analysis, as this may not be a suitable representation of the whole sample. These participants did also use a different methodology as compared to other participants, and their high bias may also be related to their methodology (see below).

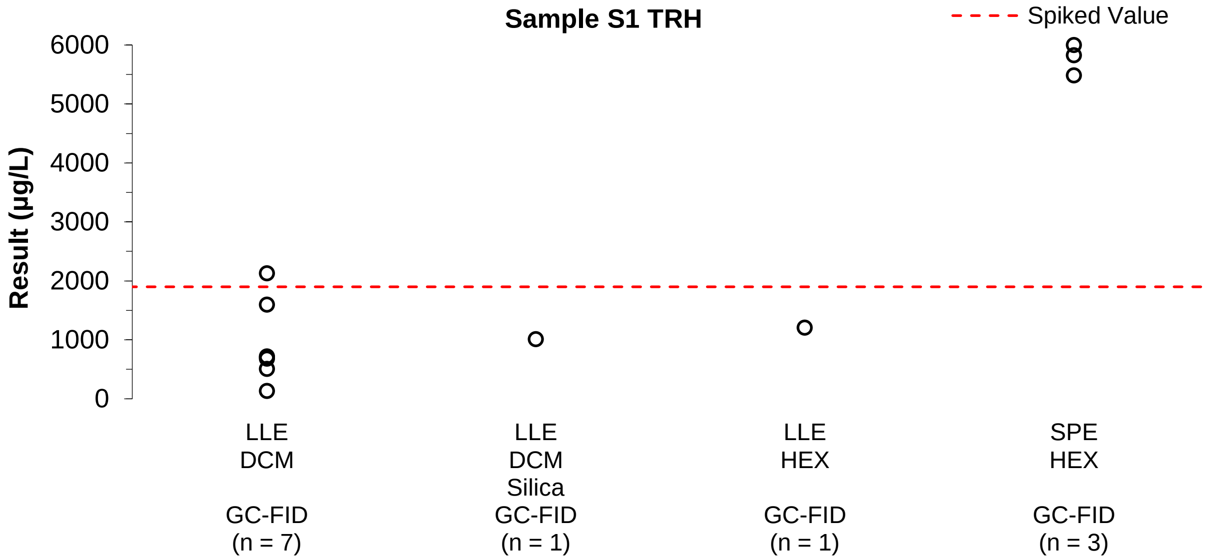


Results greater than 6000 µg/L have been plotted at 6000 µg/L.

Figure 25 Sample S1 TRH Results vs Sample Volume

Most participants reported using liquid-liquid extraction (LLE) with dichloromethane (DCM) as their extraction solvent, with one of those participants also reporting a silica clean-up step. One participant reported using LLE with hexane (HEX) instead. Three participants used solid-phase extraction (SPE) with HEX as the extraction solvent. All participants reported using gas chromatography (GC) coupled with flame ionisation detection (FID) for analysis.

A comparison of the TRH results and methodology used is presented in Figure 26. Those participants using SPE with HEX reported significantly higher TRH results.



Results greater than 6000 µg/L have been plotted at 6000 µg/L.

Figure 26 Sample S1 TRH Results vs Methodology

Sample S2 BTEX

Ten participants reported taking the whole sample (42 mL) for analysis, while others reported sample test portions ranging from 5 mL to 12 mL.

A comparison of the Total BTEX results and sample volume used is presented in Figure 27.

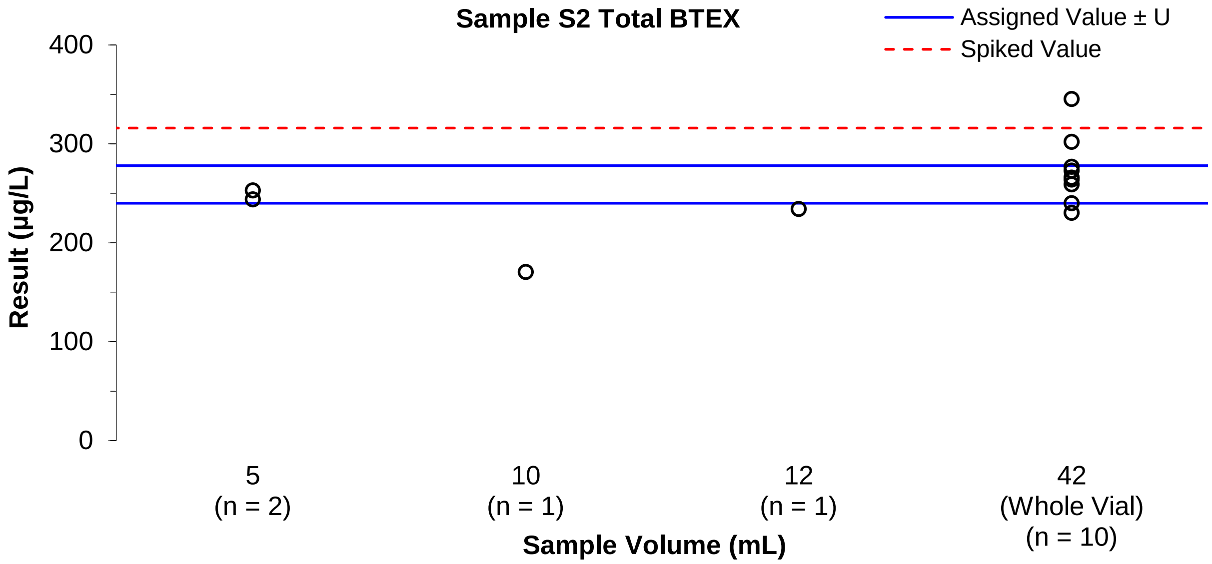


Figure 27 Sample S2 Total BTEX Results vs Sample Volume

For BTEX analysis, participants used either purge-and-trap (P&T) GC coupled to mass spectrometry (MS) or tandem mass spectrometry (MS/MS), or headspace (HS) GC-MS. Two participants reported LLE as part of their preparation. The most common methodology was P&T GC-MS.

A comparison of the Total BTEX results and methodology used is presented in Figure 28.

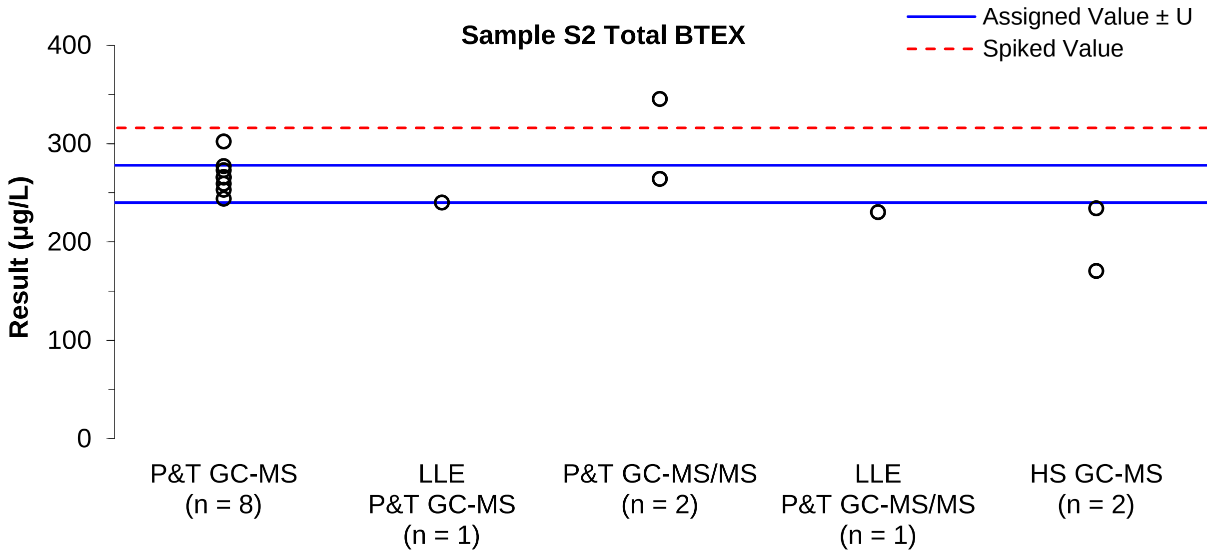


Figure 28 Sample S2 Total BTEX Results vs Methodology

Sample S3 PAHs

For this study, participants were given the option of either analysing 3 x 100 mL bottles (six participants) or 1 x 500 mL bottle (five participants), depending on which best suited their laboratory's method. Participants reported test portions ranging from 35 mL to 500 mL. No participant using 3 x 100 mL bottles used the whole bottle, whereas three participants using 1 x 500 mL bottle used the whole bottle.

A comparison of the PAHs *z*-scores and sample volume used is presented in Figure 29.

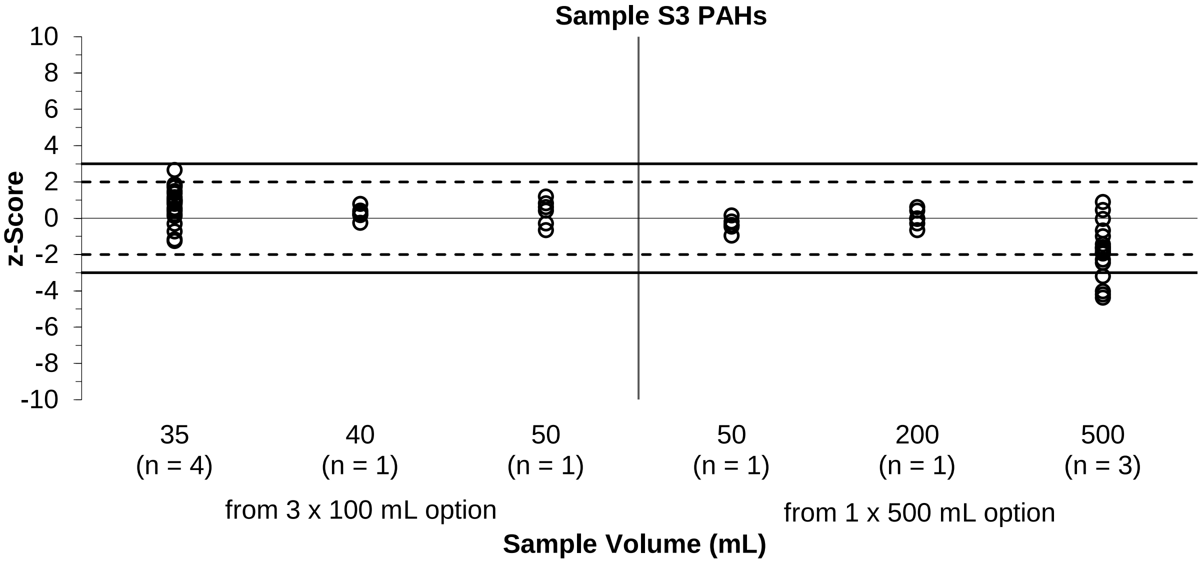


Figure 29 Sample S3 PAHs *z*-Scores vs Sample Volume

Most participants used LLE with DCM; two participants used HEX as the extraction solvent instead. Three other participants used SPE with DCM/ethyl acetate (EtOAc). No participant reported a clean-up step. All participants used GC-MS(/MS) for analysis.

A comparison of the PAHs *z*-scores and the methodology used is presented in Figure 30.

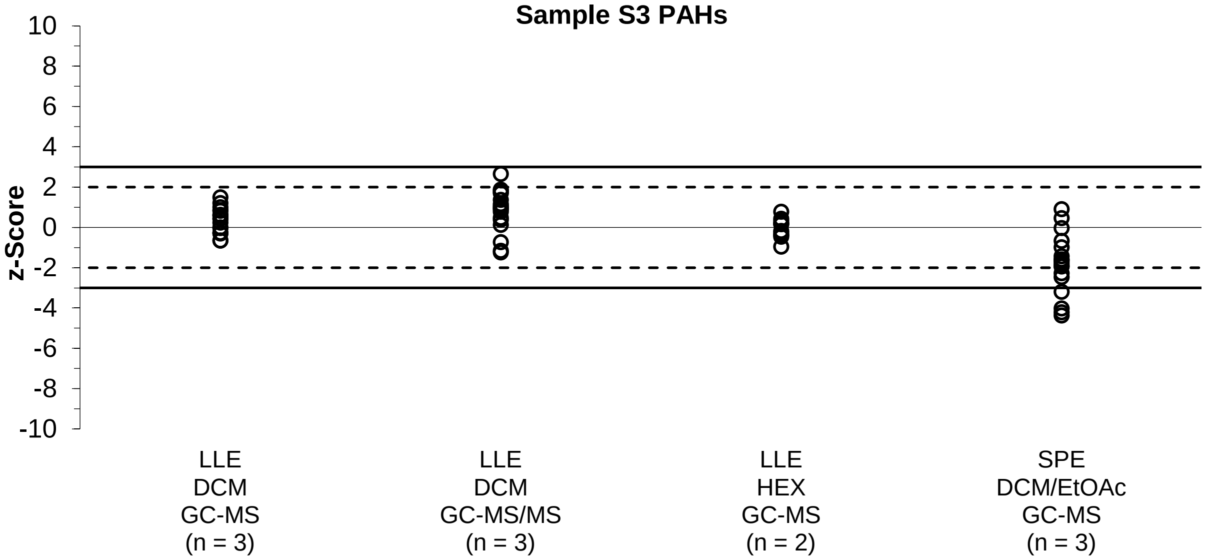


Figure 30 Samples S3 PAHs *z*-Scores vs Methodology

6.8 Certified Reference Materials

Participants were requested to report whether certified standards or matrix reference materials had been used as part of the quality assurance for the analysis.

Ten participants reported using certified standards and one participant reported using matrix reference materials. Participants reported sourcing these standards from:

- NMIA (e.g. CRM MX015 Hydrocarbon-contaminated Soil)
- o2si
- Sigma-Aldrich (e.g. CRM47505)
- Other ISO 17034 standards

It is noted that the matrix for NMIA CRM MX015 is soil, which is a different matrix to the one considered in this study (river water).

These materials may or may not meet the internationally recognised definition of a CRM:

> '***reference material***, *accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceabilities, using valid procedures*'[11]

6.9 Summary of Participants' Performance

Summaries of participants' results and performances for scored analytes in this PT study are presented in Tables 28 and 29, and Figure 31.

Table 28 Summary of Participants' Results for Sample S2 Scored Analytes*

Lab. Code	Benzene	Toluene	Ethylbenzene	Xylenes	Total BTEX
AV	53.7	105	18.4	81.7	259
SV	55.1	128	24.4	109	316
1	54.5	93.5	17.7	78.3	244
2	49.55	94.71	17.04	72.9	234.2
3	73.7	135.6	26.8	109.3	345.4
4	33	65.5	13	59	170.5
6	51.1	109.4	19.4	85.6	265.5
7	49.5	105.3	19.8	84.5	259.1
8	62	129	20	92	302
9	50.7	95.5	17.9	66.2	230.4
10	54	114	19.7	93.4	273
11	56	110	19	82	266
12	764	633	18	76	264
13	52	97	18	86	253
14	49	92	19	78	240
15	62	116	15	85	277

* All values are in µg/L. Shaded cells are results which returned a questionable or unacceptable *z*-score. AV = Assigned Value, SV = Spiked Value.

Table 29 Summary of Participants' Results for Sample S3 Scored Analytes*

Lab. Code	Anthracene	Benzo[*a*]pyrene	Chrysene	Fluoranthene	Fluorene	Pyrene
AV	3.20	1.61	2.55	2.29	1.88	1.60
SV	4.48	2.00	3.01	3.00	3.01	2.00
1	1.8	0.9	1.4	1.4	1.2	0.9
2	3.27	1.52	2.37	2.16	1.83	1.37
3	5.31	0.72	9.59	3.83	3.63	4.83
4	3.3	1.65	2.45	2.4	2.1	1.7
6	2.64	2.05	2.91	2.33	2.16	2.05
7	2.6	2.02	2.89	2.42	2.1	2.02

Lab. Code	Anthracene	Benzo[*a*]pyrene	Chrysene	Fluoranthene	Fluorene	Pyrene
8	3.18	1.06	0.88	1.62	1.4	1.01
9	3.4665	1.6648	2.4319	2.8009	2.1621	1.8062
10	2.85	2.25	2.73	2.69	2.18	1.93
11	3.42	0.84	0.94	1.95	1.43	1.21
12	3.2	1.6	2.3	2.5	1.8	1.7
13	NS	NS	NS	NS	NS	NS
14	3.6	1.9	2.3	2.5	1.8	1.7
15	3.63	1.17	1.01	1.8	1.69	1.14

* All values are in µg/L. Shaded cells are results which returned a questionable or unacceptable *z*-score. AV = Assigned Value, SV = Spiked Value.

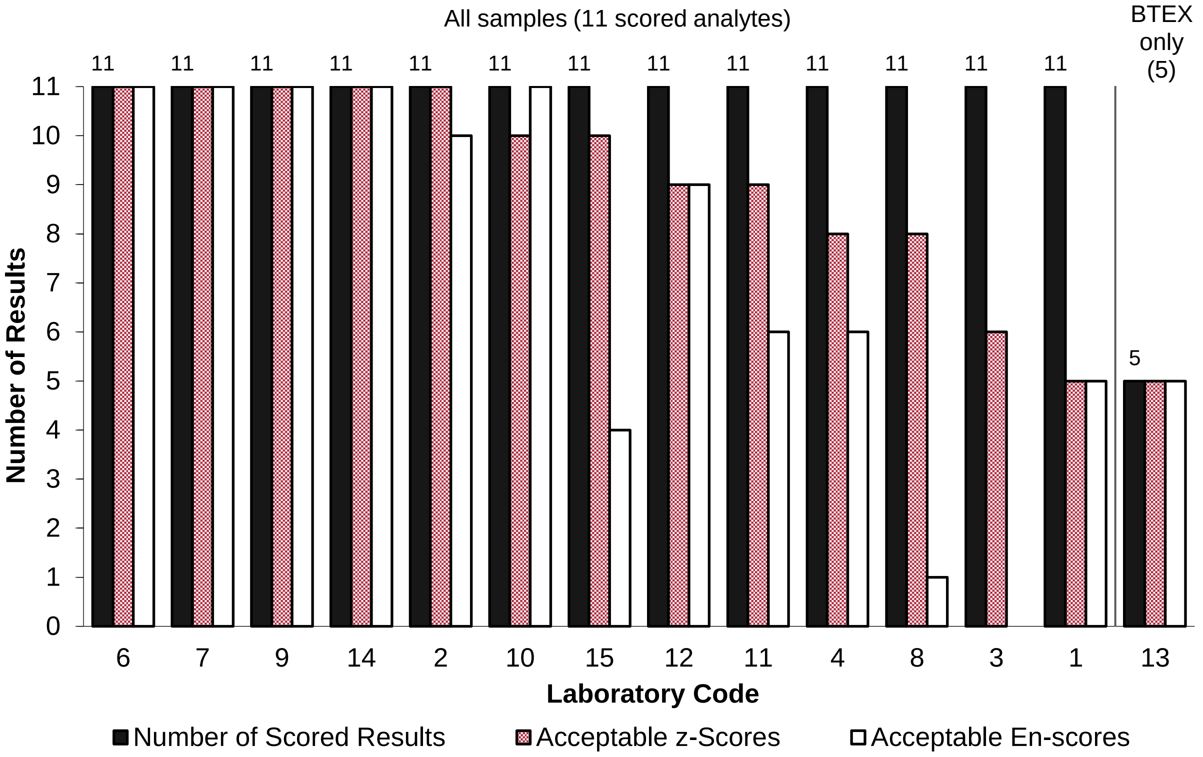


Figure 31 Summary of Participants' Performance

6.10 Comparison with Previous Studies

Over the last 10 studies (2015–2024), the proportion of acceptable scores for hydrocarbons in river water has remained fairly consistent. However, the proportion of acceptable scores in the current study AQA 24-19 is among the lowest over this period for both *z*-scores and E_n-scores. This study had a larger number of new participants.

Total BTEX

A summary of *z*-scores and E_n-scores, presented as a percentage of the total number of scores for each study, obtained by participants for total BTEX in river water over the last 10 studies (2015–2024) is presented in Figure 32. Over this period, the proportion of acceptable scores has remained high, with an average proportion of 94% and 87% for *z*-scores and E_n-scores respectively. However, the proportion of acceptable scores in this study was the second lowest over this period for *z*-scores, and the lowest over this period for E_n-scores.

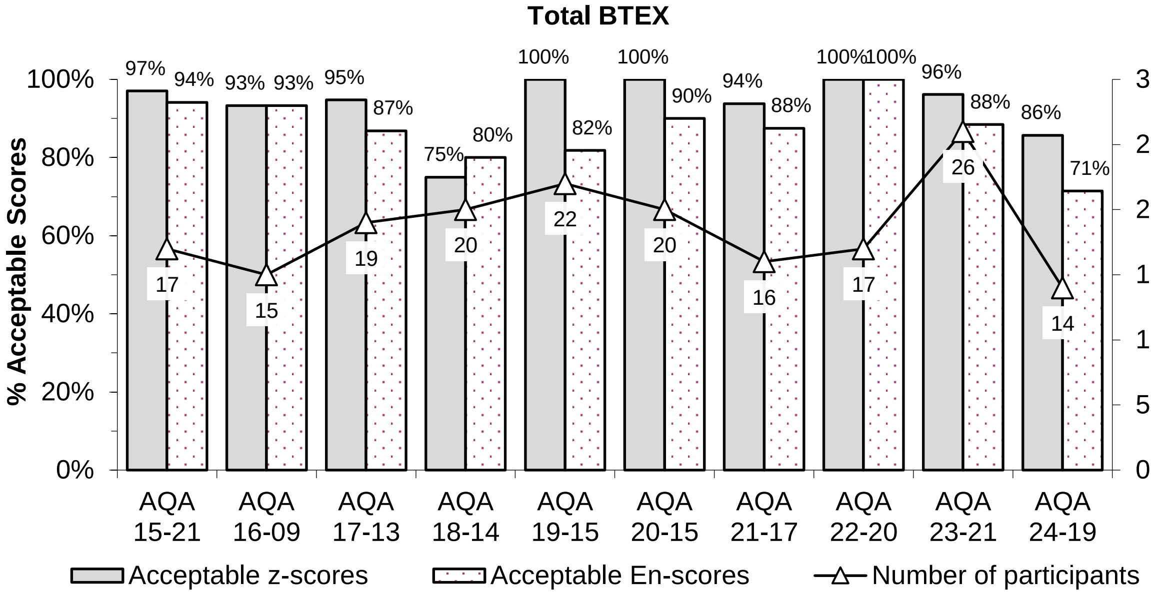


Figure 32 Summary of Acceptable Scores for Total BTEX in River Water PT Studies

PAHs

A summary of *z*-scores and E_n-scores, presented as a percentage of the total number of scores for each study, obtained by participants for PAHs in river water over the last 10 studies (2015–2024) is presented in Figure 33. Over this period, the proportion of acceptable scores has remained fairly consistent, with an average proportion of 86% and 80% for *z*-scores and E_n-scores respectively. However, the proportion of acceptable scores in this study was the lowest over this period for both *z*-scores and E_n-scores.

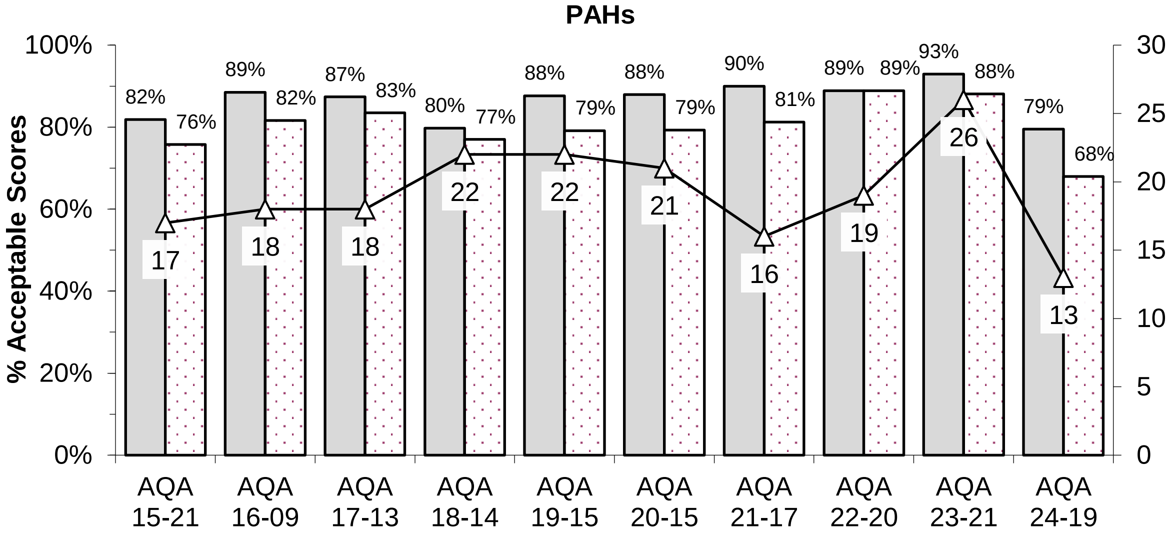
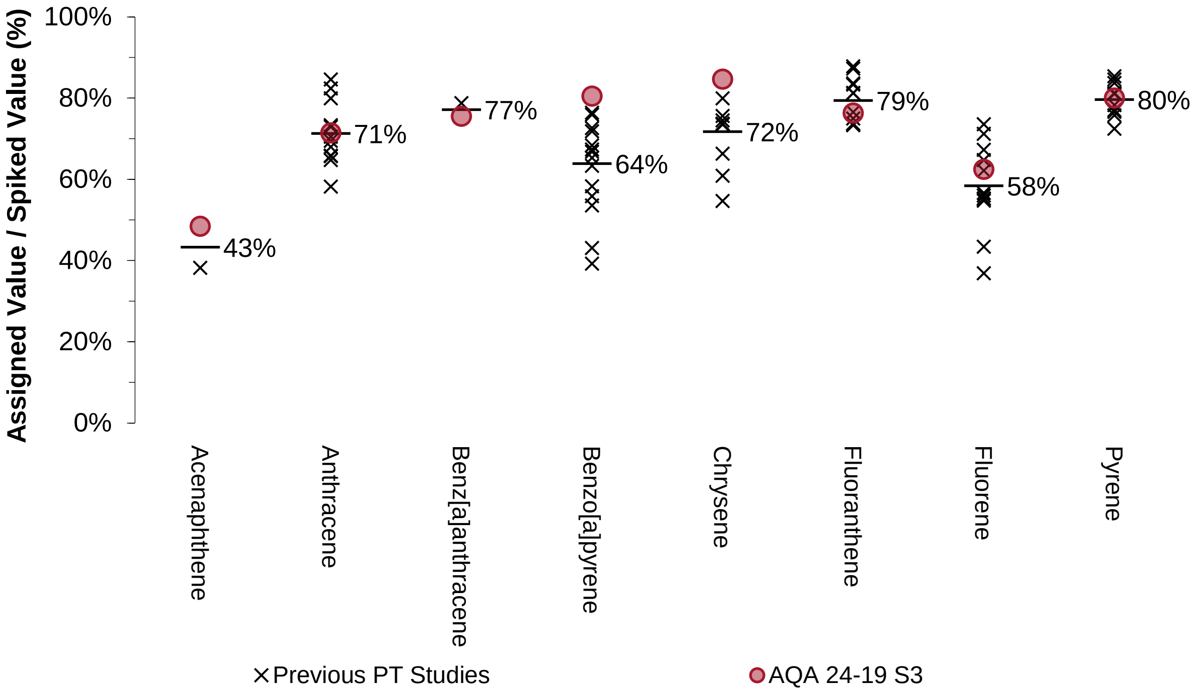


Figure 33 Summary of Acceptable Scores for PAHs in River Water PT Studies

A plot of the assigned value (or robust average where no assigned value was set), expressed as a percentage of the spiked value, for PAHs in river water since 2015 is presented in Figure 34 for Sample S3 analytes. Participants in this study returned results closer to the spiked values for benzo[*a*]pyrene and chrysene as compared to participants in previous studies.



Lines indicate the average for each PAH. Where no assigned value was set, the robust average was used instead.

Figure 34 Ratio of Assigned Value to Spiked Value for PAHs in River Water PT Studies

Measurement Uncertainties

As discussed in Section 6.2, it is a requirement of ISO/IEC 17025 that laboratories report their uncertainties if the client's instruction so requires.[9] Figure 35 presents a summary of the relative uncertainties as reported by participants over the last 10 studies (2015–2024). Over this time period, 90% of participants reported that they were accredited to ISO/IEC 17025. The vast majority (94%) of numeric results were reported with uncertainties, though this year's study had a relatively high proportion of results not being reported with uncertainties.

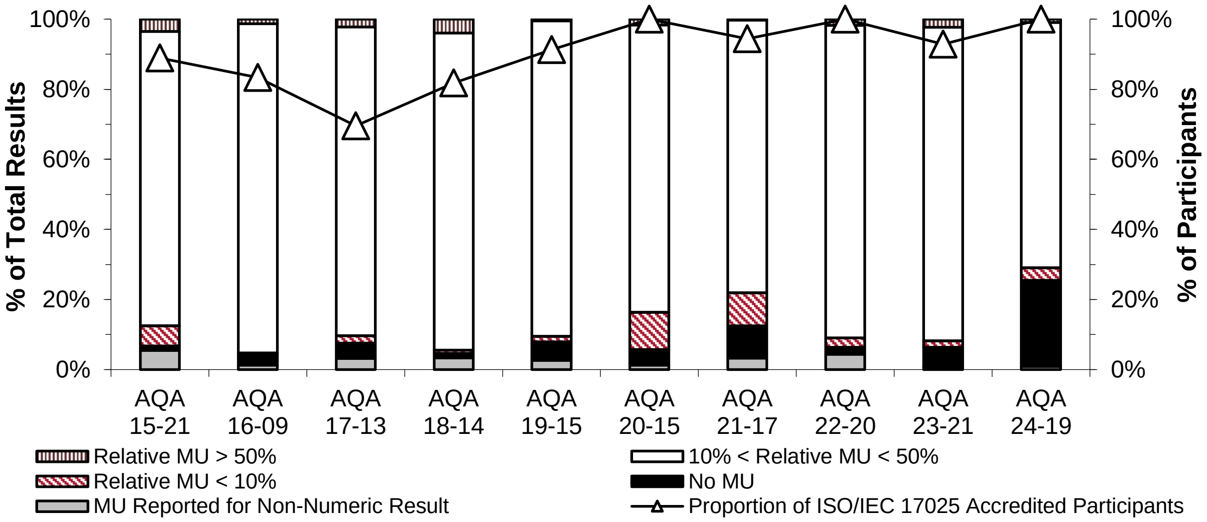


Figure 35 Summary of Participants' Relative Uncertainties for NMIA Hydrocarbons in River Water PT Studies

7 REFERENCES

Please note that for all undated references, the latest edition of the referenced document (including any amendments) applies.

[1] ISO/IEC 17043, *Conformity assessment – General requirements for the competence of proficiency testing providers.*

[2] Commonwealth of Australia, Department of Industry, Science and Resources, NMIA, 2024, *Study Protocol for Proficiency Testing*, viewed February 2025, <https://www.industry.gov.au/sites/default/files/2020-10/cpt_study_protocol.pdf>

[3] Commonwealth of Australia, Department of Industry, Science and Resources, NMIA, 2024, *Statistical Manual*, viewed February 2025, <https://www.industry.gov.au/sites/default/files/2019-07/cpt_statistical_manual.pdf>

[4] Thompson, M., Ellison, S.L.R., & Wood, R., 2006, 'The International Harmonized Protocol For The Proficiency Testing Of Analytical Chemistry Laboratories', *Pure Appl. Chem.*, vol 78, pp 145–196.

[5] National Environment Protection Council, National Environment Protection (Assessment of Site Contamination) Measure 1999 Amendment Measure 2013, viewed February 2025, <https://www.legislation.gov.au/F2008B00713/latest/text/12>

[6] La Greca, B., 1996, 'Storage Stability Study: Petrol BTEX Residues in Water', *ACSL Public Interest Project*, AGAL.

[7] ISO 13528, *Statistical methods for use in proficiency testing by interlaboratory comparison*.

[8] Thompson, M., 2000, 'Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing', *Analyst*, vol 125, pp 385–386.

[9] ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories.*

[10] Eurachem/CITAC Guide CG 4, QUAM:2012.P1, *Quantifying Uncertainty in Analytical Measurement*, 3rd edition, viewed February 2025, <http://www.eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf>

[11] JCGM 200:2012, *International vocabulary of metrology – Basic and general concepts and associated terms (VIM)*, 3rd edition.

APPENDIX 1 SAMPLE PREPARATION

A1.1 Diesel Fuel and River Water Preparation

Diesel fuel was purchased from a local retail outlet and treated to remove volatiles. Approximately 500 mL of diesel fuel was placed in an open container and sparged with nitrogen to reduce to as low as possible any analyte eluting before C_{10}.

Water was sampled from Browns Waterhole, Turramurra. The water was filtered under vacuum through an Advantec 150 mm glass fibre filter. After filtration, the water was placed in Schott bottles and autoclaved.

The bottles used for S1 and S3 were rinsed with acetone and heated to 140 °C overnight.

A1.2 Test Sample Preparation

Sample S1

A diesel spiking solution was prepared by weighing a portion of the treated diesel fuel into a 500 mL volumetric flask and making to volume with methanol. The cleaned bottles were placed in an air-conditioned room overnight. Filtered autoclaved river water (498.5 ± 0.2 g, or 500 mL at 25 °C) was weighed into the bottles. The water was pumped into the vials using a peristaltic pump. Methanol/diesel spiking solution (1180 µL) was added to each bottle using a Hamilton dispenser. The bottles were immediately capped and inverted to mix the solution. Each bottle was then labelled and shrink-wrapped.

Sample S2

Filtered autoclaved river water (41.88 ± 0.05 g, or 42 mL at 25 °C) was weighed into Agilent headspace vials. The water was pumped into the vials through a 0.2 µm Sartorius filter capsule using a peristaltic pump. A composite spike solution was prepared by adding aliquots of diesel, unleaded petrol and benzene to methanol. The composite spiking solution was made up to volume with methanol. Composite spiking solution (1.0 mL) was added to each vial using a Hamilton dispenser. Each vial was capped after spiking, and then labelled and shrink-wrapped.

Sample S3

The spiking solutions were prepared by dissolving each standard material in DCM. Diluted spiking solutions were prepared using acetone. The autoclaved water was placed in a stainless steel container. After spiking the water was stirred using a top-driven impeller stirrer for at least two hours. The samples were then dispensed into 500 mL and 100 mL amber glass bottles which were then labelled and shrink-wrapped.

Between preparation and dispatch, all samples were stored in a cool room at 4 °C.

APPENDIX 2 ASSESSMENT OF HOMOGENEITY AND STABILITY

A2.1 Homogeneity

No homogeneity testing was completed for this study prior to sample dispatch, as the samples were prepared using a process previously demonstrated to produce sufficiently homogeneous samples for similar analytes and matrices.

As results returned by participants for Sample S1 were not in good agreement with each other, three bottles were analysed to perform partial homogeneity testing for this sample. The results of this investigation did not indicate any homogeneity issues with the samples.

The results returned by participants for Samples S2 and S3 gave no reason to question these samples’ homogeneity.

A2.2 Stability

No stability testing was conducted for this study as the samples were stored and dispatched using a process previously demonstrated to produce sufficiently stable samples for similar analytes and matrices.

The results of this study also gave no reason to question the samples’ transportation stability. Comparisons of *z*-scores to days in transit are presented in Figures 36 to 42 for scored analytes (results excluded from statistics in Section 5 have also been excluded in this section).

Figure 36 S2 Total BTEX vs Transit Days

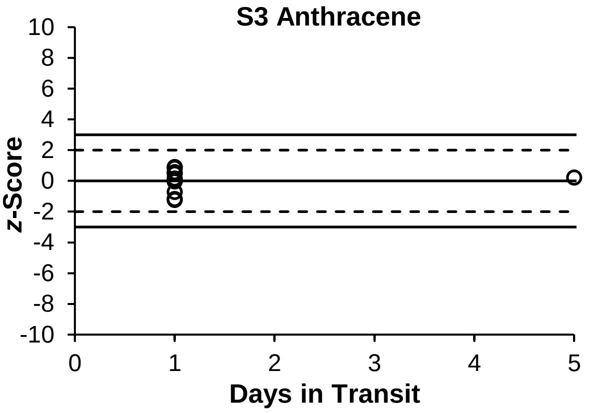


Figure 37 S3 Anthracene vs Transit Days

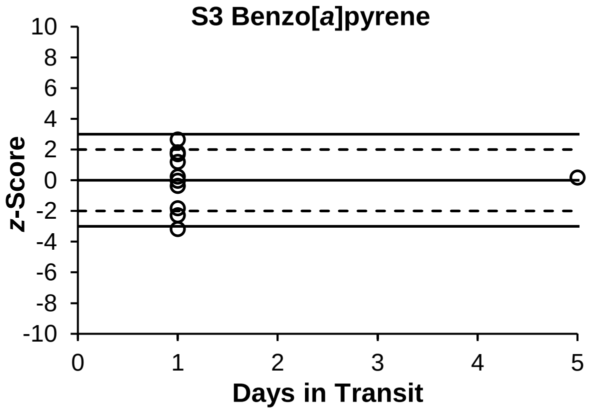


Figure 38 S3 Benzo[*a*]pyrene vs Transit Days

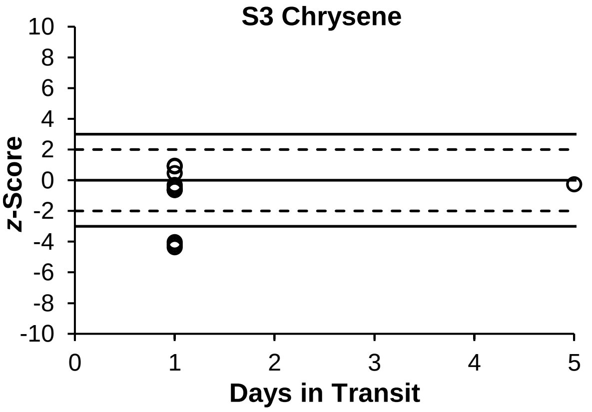


Figure 39 S3 Chrysene vs Transit Days

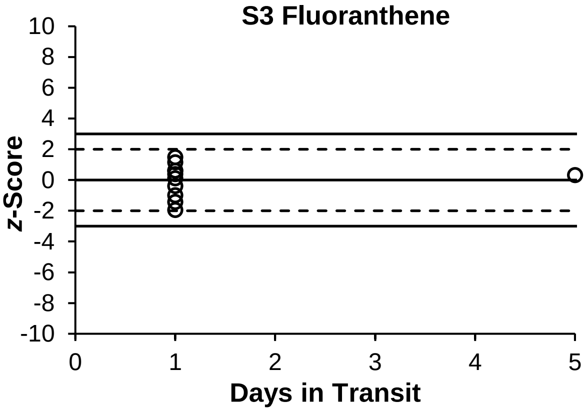


Figure 40 S3 Fluoranthene vs Transit Days

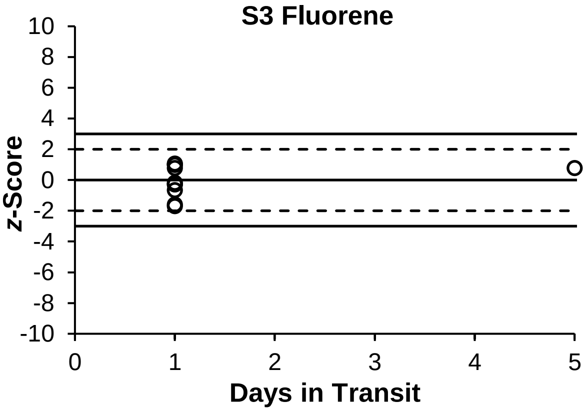


Figure 41 S3 Fluorene vs Transit Days

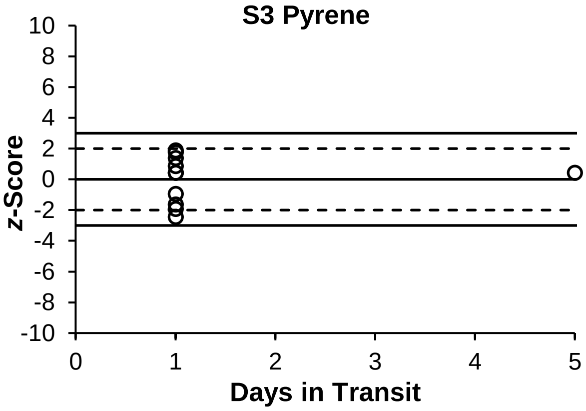


Figure 42 S3 Pyrene vs Transit Days

APPENDIX 3 ROBUST AVERAGE AND ASSOCIATED UNCERTAINTY, z-SCORE AND E$_n$-SCORE CALCULATIONS

A3.1 Robust Average and Associated Uncertainty

Robust averages were calculated using the procedure described in ISO 13528.[7] The associated uncertainties were evaluated as according to Equation 4.

$$u_{rob\ av} = 1.25 \times \frac{S_{rob\ av}}{\sqrt{p}} \qquad \textit{Equation 4}$$

where:

$u_{rob\ av}$	is the standard uncertainty of the robust average
$S_{rob\ av}$	is the standard deviation of the robust average
p	is the number of results

The expanded uncertainty ($U_{rob\ av}$) is the standard uncertainty multiplied by a coverage factor of 2 at approximately 95% confidence level.

A worked example is set out below in Table 30.

Table 30 Uncertainty of the Robust Average for Sample S2 Ethylbenzene

Number of results (p)	25
Robust Average	18.4 µg/L
$S_{rob\ av}$	1.8 µg/L
$u_{rob\ av}$	0.6 µg/L
k	2
$U_{rob\ av}$	1.2 µg/L

Therefore, the robust average for ethylbenzene in Sample S2 is 18.4 ± 1.2 µg/L.

A3.2 z-Score and E$_n$-Score Calculation

For each participant's result, a *z*-score and E_n-score are calculated according to Equations 2 and 3 respectively (Section 4).

A worked example is set out below in Table 31.

Table 31 *z*-Score and E_n-Score for Sample S2 Ethylbenzene Result Reported by Laboratory 1

Participant Result (µg/L)	Assigned Value (µg/L)	Target Standard Deviation	*z*-Score	E_n-Score
17.7 ± 4	18.4 ± 1.2	15% as PCV, or: 0.15 × 18.4 = 2.76 µg/L	$z = \frac{17.7 - 18.4}{2.76} = -0.25$	$E_n = \frac{17.7 - 18.4}{\sqrt{4^2 + 1.2^2}} = -0.17$

APPENDIX 4 PARTICIPANTS' TEST METHODS

Participants were requested to provide information about their test methods. Responses are presented in Tables 32 to 34. Some responses may be modified so that the participant cannot be identified.

Table 32 Methodology – Sample S1 TRH

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	500	Liquid-Liquid	DCM	Silica	GC-FID	In house
2	40	Liquid-Liquid	Hexane		GC-FID	NEPM
3	10	Liquid-Liquid	DCM	None	GC-FID	USEPA 8260
4	40	Liquid-Liquid	Hexane	none	GC-FID	
6	35	Liquid-Liquid	DCM	None	GC-FID	USEPA 8015
7	35	Liquid-Liquid	DCM	None	GC-FID	USEPA 8260
8	10	SPE	Hexane	None	GC-FID	EPA 8015D
9	35	Liquid-Liquid	DCM	None	GC-FID	In house
10	35mL	Liquid-Liquid	DCM		GC-FID	In house
11	10	SPE	Hexane	None	GC-FID	EPA 8015D
12	200	Liquid-Liquid	DCM	None	GC-FID	
13	250	Liquid-Liquid	DCM	None	GC-FID	In house
14	50	Liquid-Liquid	DCM	None	GC-FID	
15	10	SPE	Hexane	None	GC-FID	EPA 8015D

Table 33 Methodology – Sample S2 BTEX

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	5	None		None	P&T GC-MS	USEPA 8260
2	12	Headspace			Headspace GC-MS	USEPA 8260
3	42 mL			None	P&T GC-MS/MS	USEPA 5030

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
4	10	HS-GCMS	none	none	Headspace GC-MS	
6	40	Purge and Trap	N/A	None	P&T GC-MS	USEPA 8260
7	43	Purge and trap	None	None	P&T GC-MS	USEPA 8260
8	40	Direct Injection	None	None	P&T GC-MS	EPA 524.3
9	40	Liquid-Liquid		None	P&T GC-MS/MS	In house
10	43mL	none	none	none	P&T GC-MS	USEPA 8260
11	40	Direct Injection	None	None	P&T GC-MS	EPA 524.3
12	42	Purge and Trap	None	None	P&T GC-MS/MS	
13	5	Purge & Trap	NA	None	P&T GC-MS	USEPA 524.2
14	44	Liquid-Liquid	Methanol	None	P&T GC-MS	
15	40	Direct Injection	None	None	P&T GC-MS	EPA 524.3

Table 34 Methodology – Sample S3 PAHs

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	250	SPE	DCM/EtOAc	None	GC-MS	USEPA 8270
2	50	Liquid-Liquid	Hexane		GC-MS	USEPA 8270
3	10	Liquid-Liquid	DCM	None	GC-MS/MS	USEPA 8270
4	40	Liquid-Liquid	Hexane	none	GC-MS	
6	35	Liquid-Liquid	DCM	None	GC-MS/MS	USEPA 8270
7	35	Liquid-Liquid	DCM	None	GC-MS/MS	USEPA 8260
8	500	SPE	DCM/EtOAc	None	GC-MS	EPA 525.3
9	35	Liquid-Liquid	DCM	None	GC-MS	In house
10	35mL	Liquid-Liquid	DCM	None	GC-MS/MS	In house

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
11	500	SPE	DCM/EtOAc	None	GC-MS	EPA 525.3
12	200	Liquid-Liquid	DCM	None	GC-MS	
13	NS					
14	50	Liquid-Liquid	DCM	None	GC-MS	
15	500	SPE	DCM/EtOAc	None	GC-MS	EPA 525.3

APPENDIX 5 ACRONYMS AND ABBREVIATIONS

AV	Assigned Value
BTEX	Benzene, Toluene, Ethylbenzene, Xylenes
CITAC	Cooperation on International Traceability in Analytical Chemistry
CRM	Certified Reference Material
CV	Coefficient of Variation
DCM	Dichloromethane
EtOAc	Ethyl Acetate
FID	Flame Ionisation Detection
GC	Gas Chromatography
GUM	Guide to the expression of Uncertainty in Measurement
HEX	Hexane
HS	Headspace
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
k	Coverage factor
LLE	Liquid-Liquid Extraction
LOR	Limit Of Reporting
Max	Maximum
Md	Median
Min	Minimum
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MU	Measurement Uncertainty
N	Number of numeric results
NATA	National Association of Testing Authorities, Australia
NEPM	National Environmental Protection Measure
NMIA	National Measurement Institute, Australia
NR	Not Reported
NS	Not Supplied
NT	Not Tested
P&T	Purge and Trap
PAH	Polycyclic Aromatic Hydrocarbon
PCV	Performance Coefficient of Variation
PT	Proficiency Testing

RA	Robust Average
RM	Reference Material
SD	Standard Deviation
SI	International System of Units
SPE	Solid Phase Extraction
SS	Spiked Samples
SV	Spiked Value (or formulated concentration of a PT sample)
TRH	Total Recoverable Hydrocarbons
USEPA	United States Environmental Protection Agency

END OF REPORT