Australian Government
Department of Industry, Science and Resources

National Measurement Institute

Proficiency Test Final Report AQA 25-08A Heroin

December 2025

Our purpose is to help the government build a better future for all Australians through enabling a productive, resilient and sustainable economy, enriched by science and technology.

measurement.gov.au

This report replaces AQA 25-08.

REVISION HISTORY

Date	Report Number	Reason for Review
November 2025	AQA 25-08	Final Report – Original issue.
December 2025	AQA 25-08A	Corrected spelling errors for cutting agents reported by Laboratory 29 (Table 10).

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.

Sofia Racomelara

Jenny Xu

Lesley Johnston

Raluca Iavetz

Manager, Chemical Reference Values

105 Delhi Rd, North Ryde, NSW 2113, Australia

Phone: +61 2 9449 0234

Email: proficiency@measurement.gov.au

Accredited for compliance with ISO/IEC 17043

TABLE OF CONTENTS

SUMMARY

This report presents the results of proficiency study AQA 25-08 Heroin. The sample set consisted of three samples each containing heroin. A total of 32 laboratories received the samples, and 33 sets of results were reported. One laboratory submitted two sets of results, generated by different analysts.

Samples were prepared at the NMIA Sydney laboratory using heroin hydrochloride samples supplied by the Australian Federal Police.

The assigned values for duplicate Samples S1 and S2 were the robust average of the means of the results reported by each participant for these samples. The assigned value for Sample S3 was the robust average of participants' results. The associated uncertainties were evaluated from the robust standard deviations of the participants' results.

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

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

- *Assess participants' capabilities to measure heroin in samples typical of a routine seizure.*

Of 99 z-scores, 78 (79%) returned a z-score of $|z| \leq 2.0$, indicating an acceptable performance.

Of 99 E_n-scores, 81 (82%) returned a E_n-score of $|E_n| < 1.0$, indicating agreement of the participant's results with the assigned value within their respective expanded uncertainties.

Laboratories **1**, **2**, **4**, **5**, **7**, **11**, **13**, **16**, **17**, **19**, **20**, **22**, **23**, **24**, **26**, **27**, **28**, **29**, and **32** returned acceptable z-scores and E_n-scores for all three samples.

- *Develop the practical application of measurement uncertainty, and provide participants with information that will assist uncertainty evaluations.*

Of 99 numeric results, 93 (94%) were reported with an associated measurement uncertainty. Participants used a variety of methods to evaluate their measurement uncertainty. These methods produced relative uncertainties ranging from 2% to 45% of the reported results.

- *Assess participants' ability to identify cutting agents commonly found in controlled drug preparation.*

Duplicate Samples S1 and S2 was cut with quinine and Sample S3 was cut with glucose.

Of the 33 sets of submitted results, 26 participants (79%) correctly reported on the identity of at least one cutting agent in the samples.

Laboratories **1**, **8**, **11**, **24**, **30**, and **31** correctly identified all cutting agents in this study.

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

The samples produced for this study are homogeneous and are well characterised. Surplus of these samples is available for purchase and can be used for quality control and for 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:

- controlled drug assay, drugs in wipes, and clandestine laboratory;
- per- and polyfluoroalkyl substances in soil, biosolids, water, biota, food, and consumables;
- hydrocarbons, phenols and other organic compounds in soil and water;
- pesticide residues in soil, water, fruit, vegetables, and herbs;
- metals in soil, water, food, filters, and paint;
- nutrients, anions and physical tests in water and soil; and
- chlorophyll a in water.

1.2 Study Aims

The aims of the study were to:

- assess participants' capabilities to measure heroin in samples typical of a routine seizure;
- develop the practical application of measurement uncertainty, and provide participants with information that will assist uncertainty evaluations;
- assess participants' ability to identify cutting agents commonly found in controlled drug preparation; 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 PT studies 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 and The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories.[1,4]

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

2 STUDY INFORMATION

2.1 Study Timetable

The timetable of the study was:

Invitations sent	22/04/2025
Samples sent	28/07/2025
Results due	6/10/2025
Interim Report	9/10/2025
Preliminary Report	13/10/2025

Due to sample delivery delays to some international participants, the project timeline was extended.

2.2 Participation and Laboratory Code

Thirty-two laboratories registered to participate, with one laboratory requesting two sets of samples to be analysed independently by different analysts (total of 33 participants). Each participant was assigned a confidential laboratory code number for this study. All participants submitted results.

2.3 Test Material Specification

Three test samples were prepared in June 2025. The starting material was two different baches of heroin hydrochloride (approximately 80% heroin base (m/m) and 72% heroin base (m/m)) supplied by the Australian Federal Police.

Quinine purchased from Sigma-Aldrich and glucose purchased from a local pharmacy were used as cutting agents. Duplicate Samples S1 and S2 were prepared using the higher purity heroin and cut with quinine. Sample S3 was prepared using the lower purity heroin and cut with glucose.

The heroin hydrochloride was ground and sieved through a 180 μm sieve. The cutting agents were processed similarly. Test samples were prepared by mixing a known mass of sieved drug with known amounts of sieved cutting agent in a tumbler overnight. Portions of 150 mg of each of the test samples were weighed into labelled glass vials.

Samples S1 and S2 were prepared to contain approximately 61.0% heroin base (m/m).

Sample S3 was prepared to contain approximately 28.2% heroin base (m/m).

2.4 Test Sample Homogeneity and Stability

The preparation of homogeneous test samples is an important part of a PT study. Given the small (< 150 mg) test portions normally used for controlled substance analysis, the particle size must be sufficiently small and uniformly distributed to ensure minimal influence on analytical precision.

All samples were prepared using the same procedure as previous controlled drug PT studies, which has been demonstrated to produce sufficiently homogeneous samples. Results returned by the participants gave no reason to question the homogeneity of the test samples.

To assess the stability of the samples, results returned by participants were compared to the dates of analysis (Section 6.7). The results gave no reason to question the samples' stability.

2.5 Sample Dispatch and Receipt

A set of three test samples, with each sample containing approximately 150 mg of test material, was dispatched to the majority of international participants on 3 July 2025. For Australian participants and remaining international participants, samples were dispatched on 28 July 2025. Dispatch was organised based on when permits were received.

The following items were also sent with the samples:

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

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

2.6 Instructions to Participants

Participants were instructed as follows:

- Analyse each sample for amount of heroin base by your routine test method. It is recommended to thoroughly mix the content of each vial before taking a test portion for analysis, and to use a minimum test portion of 20 mg.
- For each sample, report % m/m heroin as base. Report this figure as if reporting to a client.
- For each result, report an evaluation of your expanded uncertainty as % m/m heroin as base.
- Report the identity of diluent(s)/adulterant(s) in the samples if this is within your normal scope of analysis.
- Give brief details of your:
    - basis of uncertainty evaluation (e.g. uncertainty budget, repeatability precision)
    - analytical method (e.g. sample treatment, instrument type, calibration method)
    - reference standard (e.g. source, purity)
- A results spreadsheet will be emailed to you. Please complete the results spreadsheet and return by email to jenny.xu@measurement.gov.au.
- Results are to be returned by 1 September 2025.

The results due date was later changed to 6 October 2025. This was to accommodate sample delivery delays to some international participants.

2.7 Interim Report and Preliminary Report

An Interim Report was emailed to all participants on 9 October 2025.

A Preliminary Report was emailed to all participants on 13 October 2025. This report included a summary of the results reported by participants, assigned values, performance coefficients of variation, z-scores and E_n-scores for each analyte in this study. No data from the Preliminary Report has been changed in the present Final Report.

3 PARTICIPANT LABORATORY INFORMATION

3.1 Test Methods Reported by Participants

Participants were requested to provide information about their test methods. Responses received are presented in Tables 1 and 2. Some responses may be modified so that the participant cannot be identified.

Table 1 Summary of Participants' Test Methods

Lab. Code	Extraction Solvent	Internal Standard	Calib. Points	Technique	Detector	Column	Comments
1*	Acetonitrile	nil	7	HPLC	DAD	Phenomenex C8 Luna 3u Narrow Bore 100 mm x 2 mm	Wavelength: 214 nm
2	Chloroform	Octacosane	5	GC	FID	HP5	
3*	Acetonitrile, acetic acid, water	NO ISTD	4	HPLC	DAD	Poroshell 120 Ec-18	
4	acetonitrile/H20 (80/20)	External standard	3	HPLC	DAD	NH2	
5	HPLC Methanol	-	1	UPLC	DAD	Thermo Scientific Hypersil-5-ODS	Wavelength: 280nm
6	Ethanol	Propylparaben	8	UPLC	DAD	BEH Shield RP18	
7*	Absolute ethanol	Tribenzylamine	6	GC	FID	DB5. 20 m column, 0.1 um film, 0.1 mm diameter	
8	Acetonitrile	Strychnine	6	GC	FID	HP1	
9	H20 ACN	N/A	3	HPLC	DAD	Luna C18	
10	Acetonitrile/Water	None	5	HPLC	UV	Kinetex 5u C18	Wavelength: 279nm
11	water/acetonitrile/ 2.5M sulphuric acid 90:10:1	None	3	HPLC	Diode Array	Shimpack XR-ODS	Wavelength: 279 nm
12*	methanol	no	7	HPLC	DAD	Poroshell 120C18 (4.6x150 mm. 2.7 microns particle size)	
13	Chloroform: Methanol (9:1)	2,2,2 Triphenyl acetophenone	1	GC	FID	HP5	
14	ethanol:dimethylformamide (9:1)	tribenzylamine	6	GC	FID	HP1	

Lab. Code	Extraction Solvent	Internal Standard	Calib. Points	Technique	Detector	Column	Comments
15*	Methanol	N/A	4	HPLC	DAD	Hyper-sil Gold C18 Selectivity HPLC columns (Thermofisher)	Wavelength: 205nm
16	90:10 Water: Acetonitrile (0.1% TFA)	N/A	5	HPLC	PDA	XSelect CSH, C18 3.5µm, 4.6x100mm	Wavelength: 260 nm
17	Methanol	Mepivacaine	4	LC	UV	Kinetex EVO C18 2.6µm 2.1x100mm	
18	Acetonitrile/Methanol (95:5)	Pholocodine 1mg/ml	5	UPLC	PDA	ACQUITY C-18	
19	HPLC Methanol	-	1	UPLC	DAD	Thermo Scientific Hypersil-5-ODS	Wavelength: 280nm
20*	Methanol	-	5	LC	DAD	Luna Omega 3 µm PS C18 100 Å, 150 x 4.6 mm	Wavelength: 205 nm
21	Methanol	none	5	HPLC	DAD	Kinetex C-18-xb	Wavelength: 215 nm
22	water, acetonitrile and tetrabutylammonium phosphate	none	6	HPLC	UV/Vis	NH2	
23	Chloroform & methanol	2,2,2-triphenylacetophenone	1	GC	FID	DB-5	
24*	Ethanol	2,2,2-triphenylacetophenone	3	GC	FID	HP1	
25	methanol	LOXAPINE	5	HPLC	DAD	XTERRA	
26	Methanol	Methadone	4	GC	FID	Rxi-5ms	
27	Methanol	none	2	HPLC	DAD	Luna 3 µm PFP 100 Å 150x4.6 mm	
28	Ethanol	Eicosane	6	GC	FID	HP5	
29*	acetonitrile/water (86/14)	none	4	HPLC	DAD	NH2	Wavelength: 280 nm
30*	Chloroform	C24	4	GC	FID	HP Ultra-1	
31	acetonitrile / water	none	1	HPLC	UV/Vis	Kromasil	Wavelength: 280 nm
32	Chloroform	2,2,2-triphenylacetophenone (benzopinacolone (BZP))	1	GC	FID	HP-1 12m x 0.200µm, 0.33µm film thickness	
33	MeOH	Heroin-D9	7	HPLC	MS/MS	C18	

*Additional information in Table 2.

Table 2 Test Methods Additional Comments

Lab. Code	Participant Comments
1	Acetylcodeine and monoacetylmorphine were identified as components of the powders in S1, S2 and S3
3	0 ; 5 ; 20 ; 100 mg/l
7	Dilution of sample in 10 mL of iSTD (0.25 mg/mL of TBA in abs. ETOH)
12	external standard
15	Number of calibration points reflective of purities encountered in live casework.
20	The original results (in heroin HCl form) have been converted to base using a base conversion factor 0.91.
24	10mL of the following IS solution added to sample: 750mg of TPAP dissolved in 30mL DCM, made up to 1L in ethanol.
29	Eluent acetonitrile/water (86/14)+2.25 ml pic A/litre
30	Different IS than usual, due to Quinine disrupting C28 in our method

3.2 Reported Basis of Participants' Measurement Uncertainty Evaluations

Participants were requested to provide information about their basis of measurement uncertainty (MU). Responses received are presented in Tables 3 and 4. Some responses may be modified so that the participant cannot be identified.

Table 3 Reported 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 evaluations of the method and laboratory bias k = 2	Control samples - RM / Ex PT Sample Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Matrix effects Standard purity	SWGDRUG Supplemental Document SD-4: Measurement Uncertainty for Purity Determinations
2	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - RM / Ex PT Sample Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Matrix effects Standard purity	Nordtest Report TR537
3	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - SS	Recoveries of SS	ISO 5725-2 years and ISO/TS 21748
4	Top Down - reproducibility (standard deviation) from PT studies used directly Coverage factor not reported	Control samples - CRM Duplicate analysis	Instrument calibration	Eurachem/CITAC Guide
5	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - SS Duplicate analysis	Laboratory bias from PT studies Recoveries of SS	Eurachem/CITAC Guide
6	Coverage factor not reported			
7**	Black Box k = 2	Duplicate analysis	Laboratory bias from PT studies	Eurachem/CITAC Guide
8	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - CRM Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Laboratory bias from PT studies	ISO/GUM
9	Top Down - precision and evaluations of the method and laboratory bias Coverage factor not reported	Control samples	Recoveries of SS	ISO/GUM
10	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - CRM Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Matrix effects Standard purity	Eurachem/CITAC Guide

Lab. Code	Approach to Evaluating MU	Information Sources for MU Evaluation*		Guide Document for Evaluating MU
		Precision	Method Bias	
11**	Standard deviation of replicate analyses multiplied by 2 or 3 k = 3	Control samples - CRM Duplicate analysis	Instrument calibration Standard purity	ISO/GUM
12	Top Down - precision and evaluations of the method and laboratory bias Coverage factor not reported	Control samples - CRM Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Laboratory bias from PT studies Standard purity	
13	Top Down - precision and evaluations of the method and laboratory bias k = 3	Control samples - RM / Ex PT Sample	Matrix effects Standard purity	Eurachem/CITAC Guide
14	Top Down - precision and evaluations of the method and laboratory bias Coverage factor not reported	Control samples - RM / Ex PT Sample	Standard purity	
15**	Top Down - precision and evaluations of the method and laboratory bias k = 2 (reported as 0.95)	Control samples - CRM Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Standard purity	ISO/GUM
16**	Standard deviation of replicate analyses multiplied by 2 or 3 Coverage factor not reported	Control samples - SS	Instrument calibration Masses and volumes Recoveries of SS	ISO/GUM
17	Top Down - precision and evaluations of the method and laboratory bias k = 1	Control samples - RM / Ex PT Sample Duplicate analysis		EA-04/16, 'EA guidelines on the expression of uncertainty in quantitative testing'.
18	Top Down - precision and evaluations of the method and laboratory bias k = 3	Control samples - CRM Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes	NATA GAG Estimating and Reporting Measurement Uncertainty of Chemical Test Results
19	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - SS Duplicate analysis	Laboratory bias from PT studies Recoveries of SS	Eurachem/CITAC Guide
20	Reproducibility, bias, sample heterogeneity k = 2	Control samples - RM / Ex PT Sample Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Standard purity	ENFSI Best Practise Manual, Eurachem/CITAC guide
21	Standard deviation of replicate analyses multiplied by 2 or 3 k = 3 (99%)	Control samples - RM / Ex PT Sample Duplicate analysis		Eurachem/CITAC Guide

Lab. Code	Approach to Evaluating MU	Information Sources for MU Evaluation*		Guide Document for Evaluating MU
		Precision	Method Bias	
22	Top Down - precision and evaluations of the method and laboratory bias Coverage factor not reported			ISO/GUM
23	Top Down - precision and evaluations of the method and laboratory bias k = 3	Duplicate analysis	Matrix effects	Eurachem/CITAC Guide
24	Bottom Up (ISO/GUM, fish bone/cause and effect diagram) k = 2	Control samples - CRM Duplicate analysis	Instrument calibration Homogeneity of sample Masses and volumes Matrix effects Recoveries of SS Standard purity	Eurachem/CITAC Guide
25	Coverage factor not reported			
26	Standard deviation of replicate analyses multiplied by 2 or 3 k = 2	Duplicate analysis	Masses and volumes	ISO/GUM
27	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples Duplicate analysis	Laboratory bias from PT studies	Nordtest Report TR537
28	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples - RM / Ex PT Sample Duplicate analysis	Matrix effects	ISO/GUM
29	Top Down - precision and evaluations of the method and laboratory bias k = 2	Control samples	Laboratory bias from PT studies	ISO/GUM ISO 11352
30	Top Down - precision and evaluations of the method and laboratory bias Coverage factor not reported	Control samples - RM / Ex PT Sample Duplicate analysis		Internal document based on Eurachem/CITAC, ISO/GUM
31	Standard deviation of replicate analyses multiplied by 2 or 3 Coverage factor not reported	Control samples - RM / Ex PT Sample		ISO/GUM
32	Validation k = 2			
33	Coverage factor not reported	Control samples - SS Duplicate analysis	Instrument calibration Matrix effects Standard purity	

*CRM = Certified Reference Material; RM = Reference Material; SS = Spiked Samples.
**Additional information in Table 4.

Table 4 Uncertainty Evaluation Additional Comments

Lab. Code	Participant Comments
7	Precision calculated by analysis of seized samples in laboratory reproducibility conditions.
11	MuM determined from multiple injections of reference material. 3x(Std Dev/mean)x100.
15	Uncertainty has been calculated by creating uncertainty budgets using primary source data and data obtained from sources including certificates of analysis. Measurement of uncertainty calculated following principles listed in the UKAS document M3003. MoU expanded using K=2 for 95% coverage.
16	The reported result (in routine case samples) is defined as the mean of the individual results multiplied by the uncertainty correction factor and is rounded down to the nearest whole number (unless <1%w/w). e.g a mean result of 60.2% with an uncertainty correction value of 96.64% would give a reported result of 60.2*0.9664 = 58.2 therefore rounded down to 58.0%.

3.3 Details of Participants' Calibration Standards

Participants were requested to provide information about their calibration standards. Responses as received are presented in Table 5. Some responses may be modified so that the participant cannot be identified.

Table 5 Participant Calibration Standard

Lab. Code	Reference Standard	Purity (%)
1	NMIA	99.3
2	NMIA	99.3
3	Lipomed	99.88
4	Lipomed	86.4
5	Lipomed	99.912±0.018
6	NMIA	99.4
7	LIPOMED	99.04
8	NMI	99.3
9	BP	99.3
10	British Pharmacopoeia	99.3
11	LGC Standards	99.7
12	LIPOMED	99.879
13	CRM-Lipomed	97.34
14	LGC NMIA D752d	99.3
15	Sigma-Aldrich (Merck)	99.9
16	Lipomed	99.95
17		
18	NMI	99.3
19	Lipomed	99.912±0.018
20	Chiron	97.2
21	Chiron	99.8
22	Lipomed	99.912

Lab. Code	Reference Standard	Purity (%)
23	Cayman	98.87
24	NMI	99.3
25	LGC STANDARDS	>99.9
26	LGC	1.011
27	Lipomed (M-29-FB-1LA)	1 mg/mL
28	Lipomed	98.6
29	Lipomed	98.6438
30	Fagron	100
31	Lipomed	98.64
32	NMI	99.3 ± 0.5
33	CHIRON	99.8

3.4 Participants' Comments

Participants were invited to comment on the samples, the PT study in general and suggestions for future PT studies. Such feedback allows for the improvement of future studies. Participants' comments are presented in Table 6, along with the study coordinator's response where appropriate. Some responses may be modified so that the participant cannot be identified.

Table 6 Participants' Comments

Lab. Code	Participants' Comments	Study Coordinator's Response
13	Purity of our routine seized heroin samples received by our laboratory is around 1-5%. It will good to have one of the sample vial with low purity Heroin	Thank you for your feedback. We aim to select a range of purities to cater for the needs of different laboratories. For this study, the assigned values were 61.3% base (m/m) and 27.2% base (m/m). Lower and higher heroin level samples have been included in our previous PT studies. Occasionally, identical samples are prepared to assess laboratories' performance for blind duplicate samples. This is the first heroin PT study in seven years to include blind duplicate samples.
28	I would like to raise a concern regarding the recent proficiency test. In this round, samples 1 and 2 resulted in the same concentration, which limits the ability to properly assess analytical performance across a wider range of values. This situation had already occurred in a previous round with cocaine samples. To strengthen the value of the exercise, I kindly suggest avoiding the repetition of ranges or identical concentrations in future distributions, so that laboratories can demonstrate their capability to analyse samples under more diverse conditions.	
29	We would like to receive 3 samples of very different concentrations and if possible at 3% to check our limit of quantification.	

4 PRESENTATION OF RESULTS AND STATISTICAL ANALYSIS

4.1 Results Summary

Participant results are listed in Tables 7 to 9 with resultant 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). Bar charts of results and performance scores are presented in Figures 2 to 4.

An example chart with interpretation guide is shown in Figure 1.

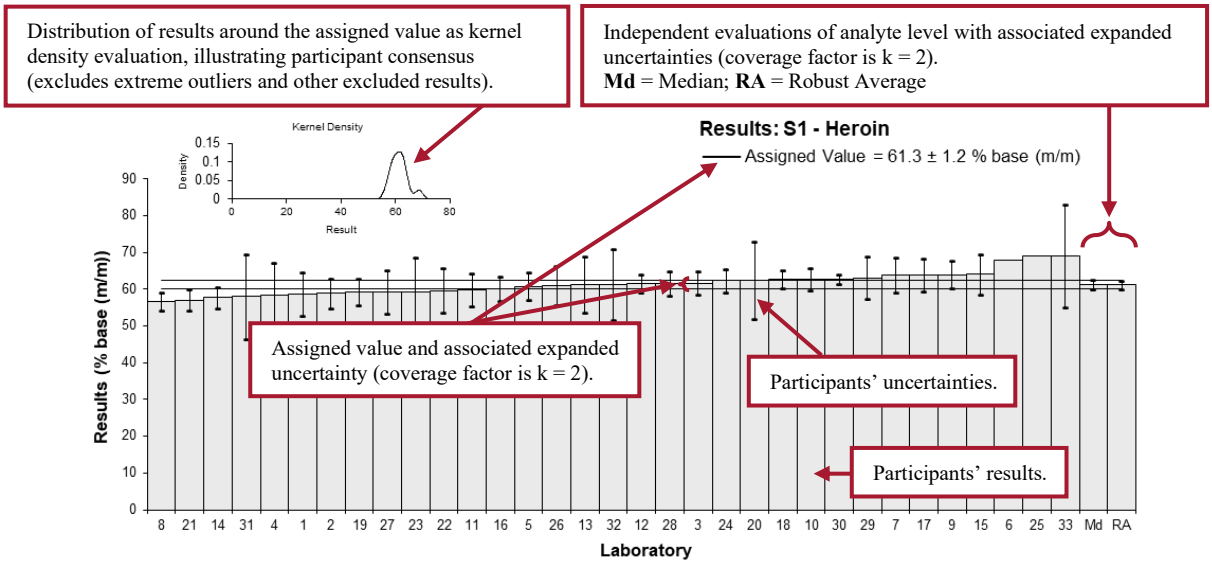


Figure 1 Guide to Presentation of Results

4.2 Outliers, Extreme Outliers and Other Excluded Results

Outliers were any result less than 50% and greater than 150% of the robust average and were removed prior to the calculation of the assigned values.[3,4] Extreme outliers (gross errors), such as those due to incorrect units, decimal placement errors, or results for a different proficiency test item, were also removed before the calculating statistics.[3]

Laboratory **9** may have switched their Sample S2 and S3 results. To avoid unfair scoring, these results were excluded from the robust average calculations as they would bias the assigned value; they were also excluded from the calculation of all summary statistics.

4.3 Assigned Value

Assigned value is defined as the 'value attributed to a particular property or characteristic of a proficiency test item'.[1] In this PT study, the property is the % heroin base (m/m) in the samples. The assigned value for the duplicate Samples S1 and S2 was determined as the robust average of the means of the results reported by each participant for these samples. The assigned value for Sample S3 was the robust averages of participants' results. The expanded uncertainties were evaluated from the associated robust SDs (Appendix 1).

4.4 Robust Average and Robust Between-Laboratory Coefficient of Variation

The robust averages and expanded MUs, and robust CVs (a measure of the variability of participants' results) were calculated as described in ISO 13528.[5]

4.5 Performance Coefficient of Variation (PCV)

The performance coefficient of variation (PCV) is a measure of the between-laboratory variation that in the judgement of the study coordinator would be expected from participants, given the levels of analytes present. The PCV is a value set by the study coordinator; it is not

calculated from the participants' results. It is based on the mass fraction of the analytes in the study and experience from previous studies, and is also supported by mathematical models such as the Thompson-Horwitz equation.[6] By setting a fixed and realistic value for the PCV, a participant's performance does not depend on other participants' performances.

4.6 Standard Deviation for Proficiency Assessment

The standard deviation for proficiency assessment (SDPA, σ) is the product of the assigned value (X) and the PCV, as presented in Equation 1. This value is used for calculation of z-scores.

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

4.7 z-Score

For each participant's 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 SDPA from Equation 1

For the absolute value of a z-score:

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

4.8 E_n-Score

The E_n-score is complementary to the z-score in assessment of laboratory performance. The 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:

- $|E_n| < 1.0$ 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.[7] Guidelines for quantifying uncertainty in analytical measurement are described in the Eurachem/CITAC Guide.[8]

5 TABLES AND FIGURES

Table 7

Sample Details

Sample No.	S1
Matrix	Powder
Analyte	Heroin
Unit	% base (m/m)

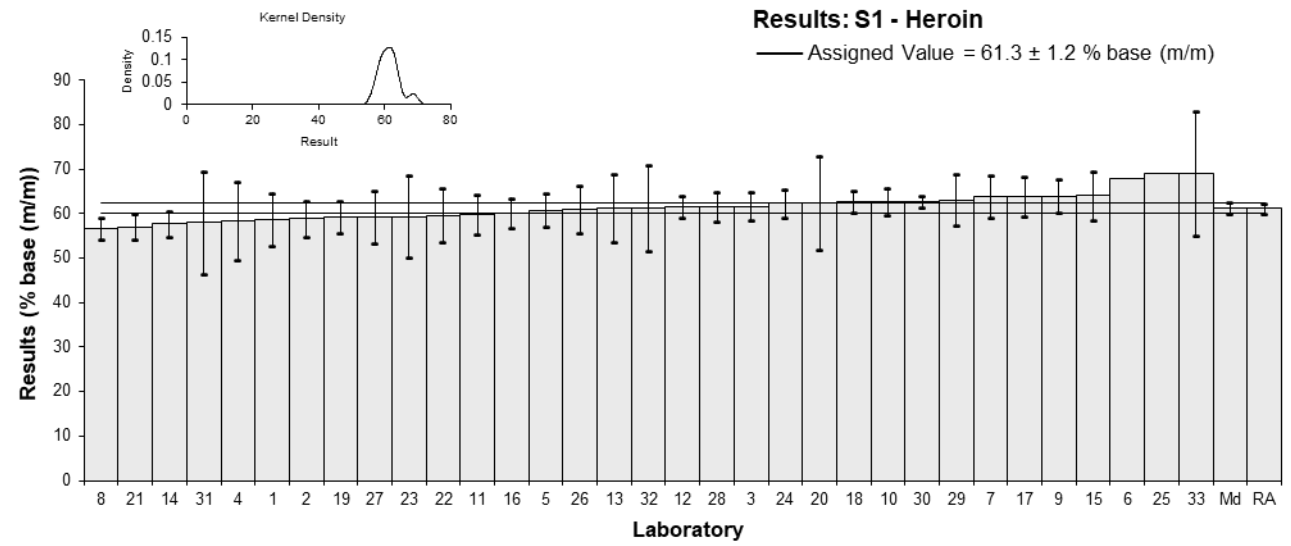
Participant Results

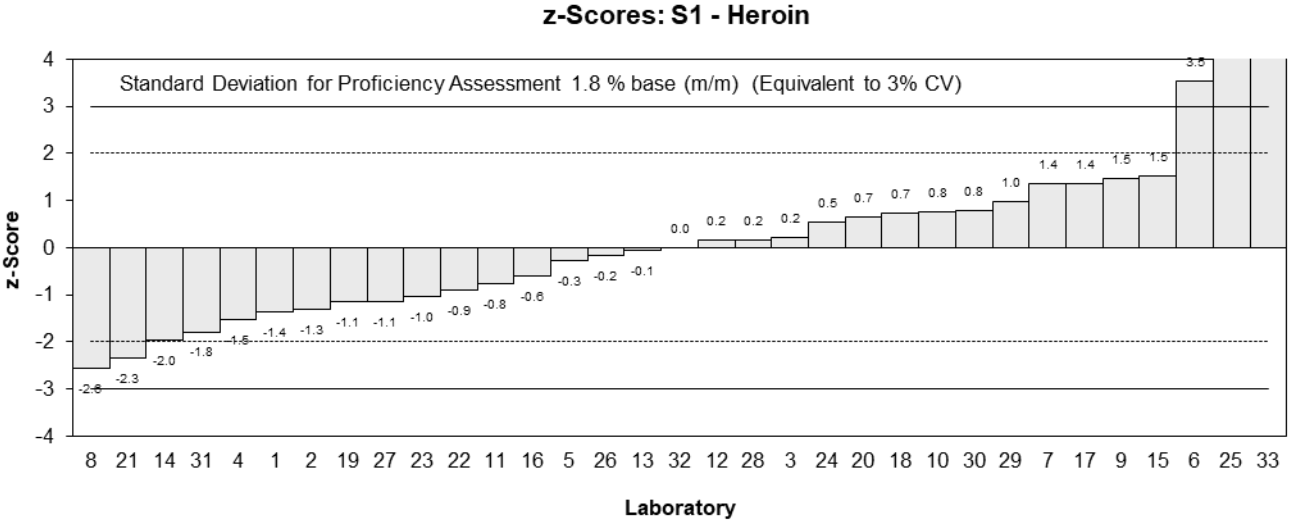
Lab. Code	Result	Uncertainty	z	E_n
1	58.8	5.9	-1.36	-0.42
2	58.9	4.1	-1.31	-0.56
3	61.7	3.3	0.22	0.11
4	58.5	8.8	-1.52	-0.32
5	60.8	3.7	-0.27	-0.13
6	67.8	NR	3.53	5.42
7	63.8	4.7	1.36	0.52
8	56.6	2.5	-2.56	-1.69
9	64	3.66	1.47	0.70
10	62.69	3.13	0.76	0.41
11	59.86	4.4	-0.78	-0.32
12	61.6	2.5	0.16	0.11
13	61.2	7.7	-0.05	-0.01
14	57.7	2.9	-1.96	-1.15
15	64.1	5.51	1.52	0.50
16	60.2	3.36	-0.60	-0.31
17	63.8	4.5	1.36	0.54
18	62.67	2.5	0.74	0.49
19	59.2	3.6	-1.14	-0.55
20	62.5	10.5	0.65	0.11
21	57	2.9	-2.34	-1.37
22	59.65	6	-0.90	-0.27
23	59.4	9.3	-1.03	-0.20
24	62.3	3.2	0.54	0.29
25	69	NR	4.19	6.42
26	61	5.22	-0.16	-0.06
27	59.2	5.9	-1.14	-0.35
28	61.6	3.3	0.16	0.09
29	63.1	5.7	0.98	0.31
30	62.75	1.19	0.79	0.86
31	58	11.6	-1.79	-0.28
32	61.3	9.6	0.00	0.00
33	69	14	4.19	0.55

Statistics

Assigned Value	61.3	1.2
Robust Average	61.2	1.2
Median	61.3	1.2
Mean	61.5	
N	33	
Max	69	
Min	56.6	
Robust SD	2.7	
Robust CV	4.5%	

The assigned value has been calculated as the robust average of the means of the results reported by each participant for duplicate Samples S1 and S2.


Kernel Density
Density
Result
Results: S1 - Heroin
Assigned Value = 61.3 ± 1.2 % base (m/m)
Results (% base (m/m))
Laboratory
8 21 14 31 4 1 2 19 27 23 22 11 16 5 26 13 32 12 28 3 24 20 18 10 30 29 7 17 9 15 6 25 33 Md RA


z-Scores: S1 - Heroin
Standard Deviation for Proficiency Assessment 1.8 % base (m/m) (Equivalent to 3% CV)
z-Score
Laboratory
8 21 14 31 4 1 2 19 27 23 22 11 16 5 26 13 32 12 28 3 24 20 18 10 30 29 7 17 9 15 6 25 33

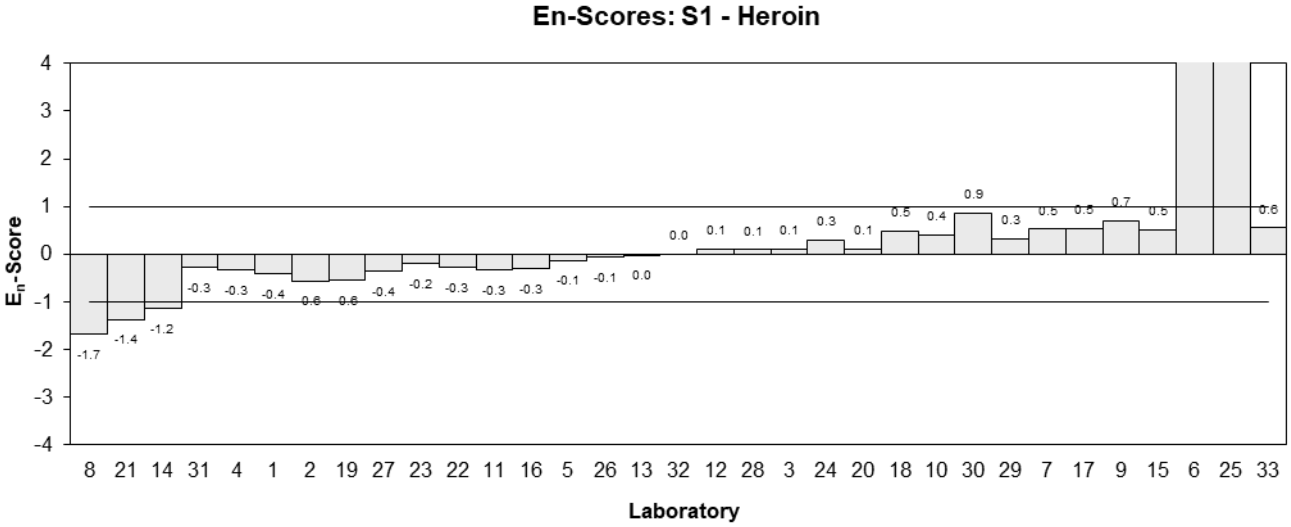

En-Scores: S1 - Heroin
En-Score
Laboratory
8 21 14 31 4 1 2 19 27 23 22 11 16 5 26 13 32 12 28 3 24 20 18 10 30 29 7 17 9 15 6 25 33

Figure 2

Table 8

Sample Details

Sample No.	S2
Matrix	Powder
Analyte	Heroin
Unit	% base (m/m)

Participant Results

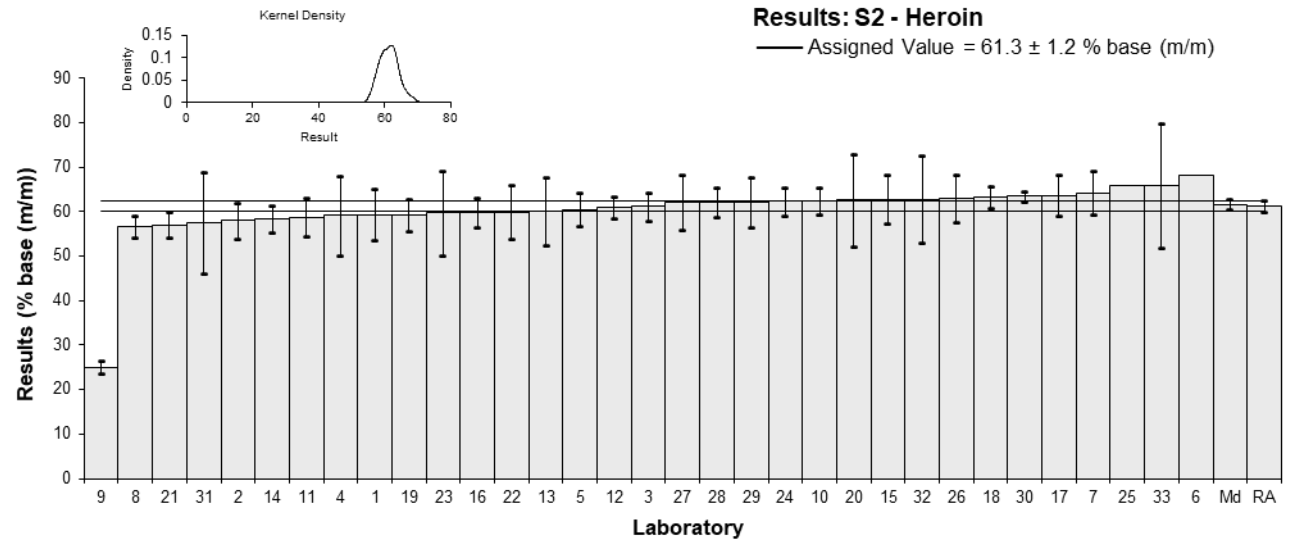
Lab. Code	Result	Uncertainty	z	E_n
1	59.4	5.9	-1.03	-0.32
2	58	4	-1.79	-0.79
3	61.2	3.2	-0.05	-0.03
4	59.2	8.9	-1.14	-0.23
5	60.5	3.7	-0.44	-0.21
6	68.3	NR	3.81	5.83
7	64.3	4.8	1.63	0.61
8	56.7	2.5	-2.50	-1.66
9**	25	1.43	-19.74	-19.45
10	62.4	3.12	0.60	0.33
11	58.72	4.32	-1.40	-0.58
12	60.9	2.4	-0.22	-0.15
13	60.1	7.6	-0.65	-0.16
14	58.4	2.9	-1.58	-0.92
15	62.77	5.51	0.80	0.26
16	59.8	3.36	-0.82	-0.42
17	63.7	4.5	1.31	0.52
18	63.37	2.5	1.13	0.75
19	59.4	3.6	-1.03	-0.50
20	62.6	10.5	0.71	0.12
21	57	2.9	-2.34	-1.37
22	59.98	6	-0.72	-0.22
23	59.7	9.4	-0.87	-0.17
24	62.3	3.2	0.54	0.29
25	66	NR	2.56	3.92
26	63	5.36	0.92	0.31
27	62.1	6.2	0.44	0.13
28	62.2	3.4	0.49	0.25
29	62.2	5.6	0.49	0.16
30	63.45	1.2	1.17	1.27
31	57.5	11.5	-2.07	-0.33
32	62.8	9.8	0.82	0.15
33	66	14	2.56	0.33

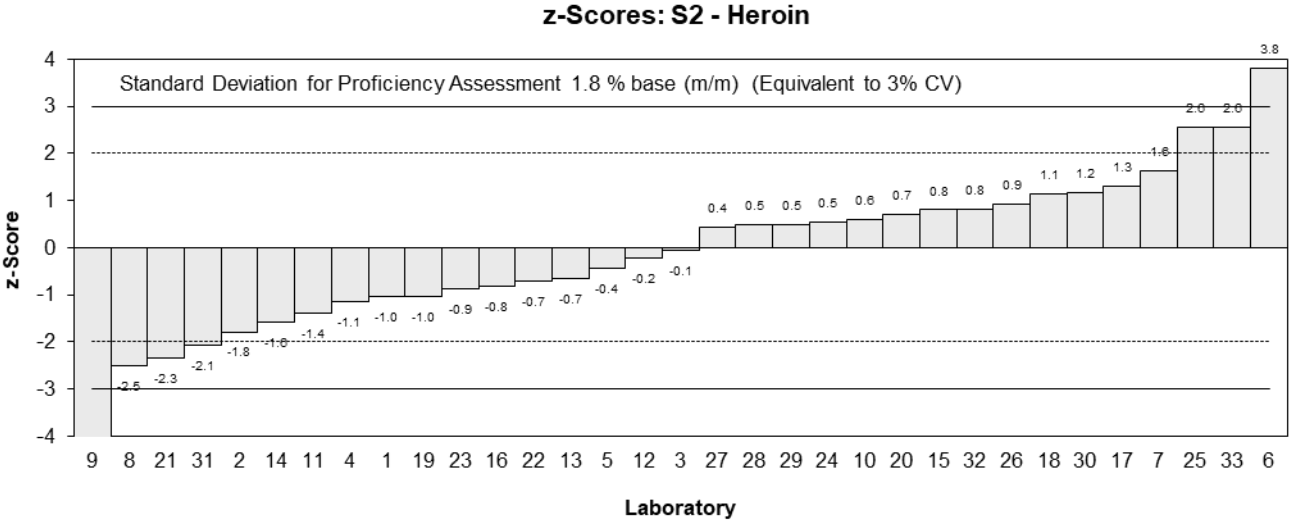
** Extreme Outlier, see Section 4.2

Statistics

Assigned Value	61.3	1.2
Robust Average	61.3	1.2
Median	61.7	1.2
Mean	61.4	
N	32	
Max	68.3	
Min	56.7	
Robust SD	2.8	
Robust CV	4.5%	

The assigned value has been calculated as the robust average of the means of the results reported by each participant for duplicate Samples S1 and S2.


Kernel Density
Density
Result
Results: S2 - Heroin
Assigned Value = 61.3 ± 1.2 % base (m/m)
Results (% base (m/m))
Laboratory


z-Scores: S2 - Heroin
Standard Deviation for Proficiency Assessment 1.8 % base (m/m) (Equivalent to 3% CV)
z-Score
Laboratory

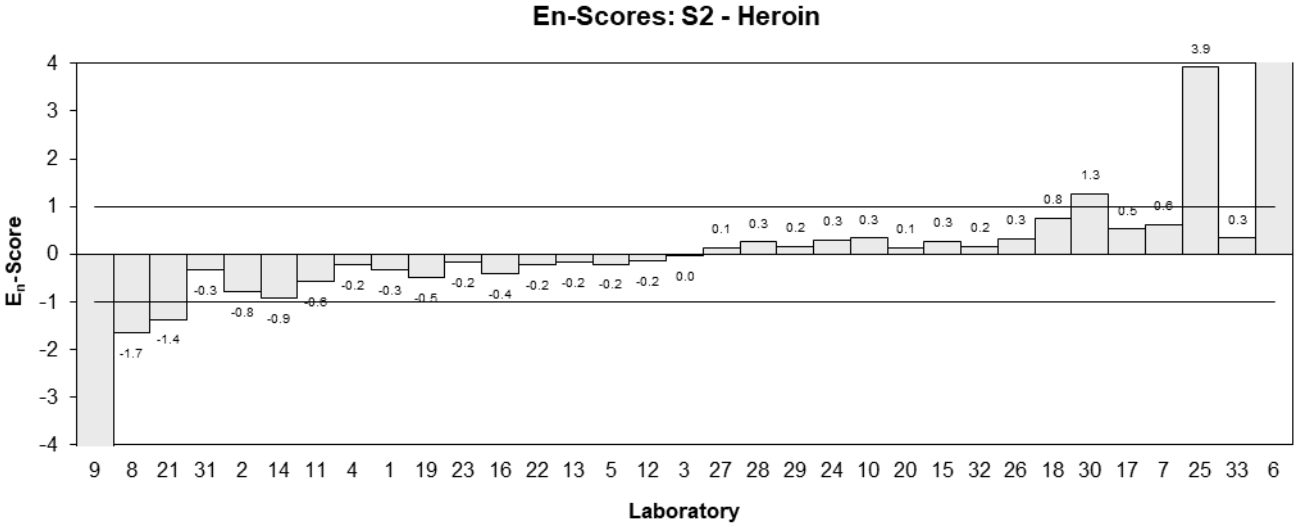

En-Scores: S2 - Heroin
En-Score
Laboratory

Figure 3

Table 9

Sample Details

Sample No.	S3
Matrix	Powder
Analyte	Heroin
Unit	% base (m/m)

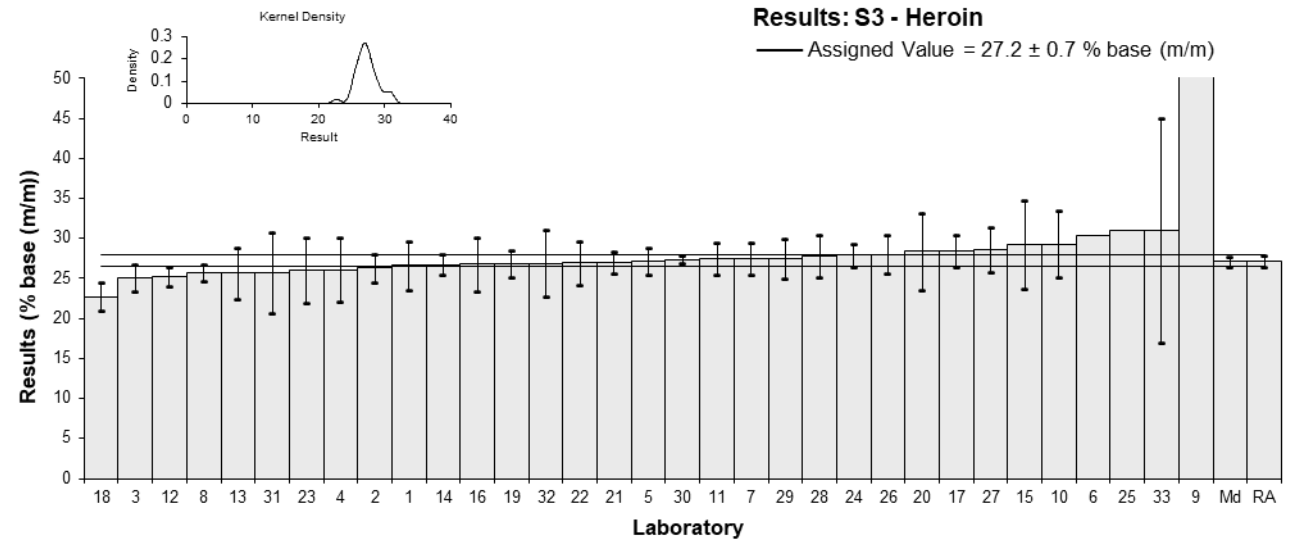
Participant Results

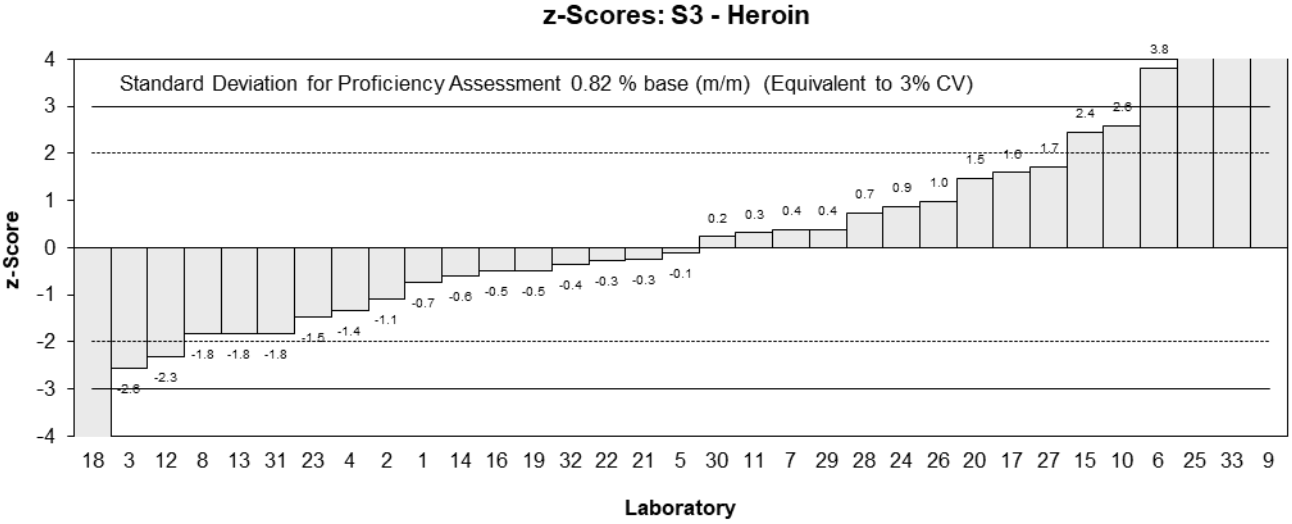
Lab. Code	Result	Uncertainty	z	E_n
1	26.6	3	-0.74	-0.19
2	26.3	1.8	-1.10	-0.47
3	25.1	1.7	-2.57	-1.14
4	26.1	4.0	-1.35	-0.27
5	27.1	1.7	-0.12	-0.05
6	30.3	NR	3.80	4.43
7	27.5	2	0.37	0.14
8	25.7	1.1	-1.84	-1.15
9**	65	3.72	46.32	9.99
10	29.31	4.1	2.59	0.51
11	27.45	2.02	0.31	0.12
12	25.3	1.2	-2.33	-1.37
13	25.7	3.2	-1.84	-0.46
14	26.7	1.3	-0.61	-0.34
15	29.19	5.51	2.44	0.36
16	26.8	3.36	-0.49	-0.12
17	28.5	2.0	1.59	0.61
18	22.74	1.7	-5.47	-2.43
19	26.8	1.7	-0.49	-0.22
20	28.4	4.8	1.47	0.25
21	27	1.4	-0.25	-0.13
22	26.97	2.7	-0.28	-0.08
23	26	4.1	-1.47	-0.29
24	27.9	1.4	0.86	0.45
25	31	NR	4.66	5.43
26	28	2.42	0.98	0.32
27	28.6	2.8	1.72	0.49
28	27.8	2.7	0.74	0.22
29	27.5	2.5	0.37	0.12
30	27.39	0.52	0.23	0.22
31	25.7	5.1	-1.84	-0.29
32	26.9	4.2	-0.37	-0.07
33	31	14	4.66	0.27

** Extreme Outlier, see Section 4.2

Statistics

Assigned Value	27.2	0.7
Robust Average	27.2	0.7
Median	27.1	0.6
Mean	27.3	
N	32	
Max	31	
Min	22.74	
Robust SD	1.5	
Robust CV	5.6%	


Kernel Density
Density
Result
Results: S3 - Heroin
Assigned Value = 27.2 ± 0.7 % base (m/m)
Results (% base (m/m))
Laboratory
18 3 12 8 13 31 23 4 2 1 14 16 19 32 22 21 5 30 11 7 29 28 24 26 20 17 27 15 10 6 25 33 9 Md RA


z-Scores: S3 - Heroin
Standard Deviation for Proficiency Assessment 0.82 % base (m/m) (Equivalent to 3% CV)
z-Score
Laboratory
18 3 12 8 13 31 23 4 2 1 14 16 19 32 22 21 5 30 11 7 29 28 24 26 20 17 27 15 10 6 25 33 9

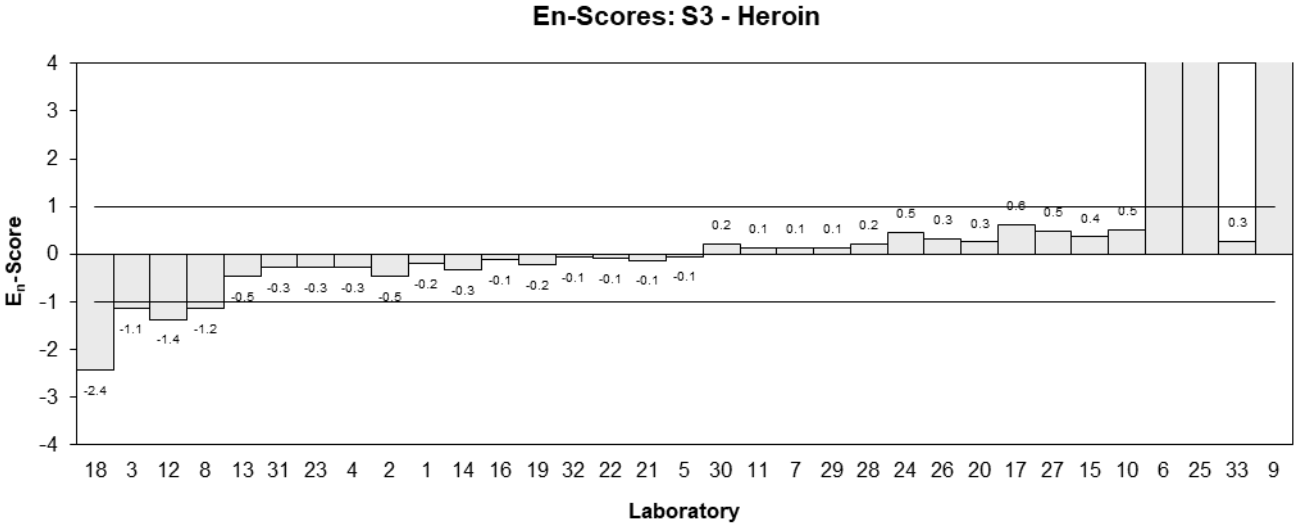

En-Scores: S3 - Heroin
En-Score
Laboratory
18 3 12 8 13 31 23 4 2 1 14 16 19 32 22 21 5 30 11 7 29 28 24 26 20 17 27 15 10 6 25 33 9

Figure 4

Table 10 Participants' Identification of Cutting Agents*

Lab. Code	Cutting Agents		
	S1	S2	S3
Preparation	Quinine		Glucose
1	Quinine	Quinine	Glucose
2	quinine	quinine	N/A
3	Quinine	Quinine	Glucose, lactose
4	Quinine	Quinine	Various sugars
5	-	-	-
6	Quinine	Quinine	
7			
8	Quinine	Quinine	Glucose
9	QUININE	Dextrose	QUININE (indicated but not confirmed)
10	Acetyl Codeine	Acetyl Codeine	Acetyl codeine, Glucose (indicated)
11	Quinine	Quinine	Glucose
12	quinine	quinine	-
13	Quinine	Quinine	Not detected
14	Quinine	Quinine	Sugars
15	N/A	N/A	N/A
16	Quinine	None	None
17	Quinine	Quinine	
18	Quinine	Quinine	
19	-	-	-
20			
21	MAM, Acetylcodeine, Quinine	MAM, Acetylcodeine, Quinine	MAM, Acetylcodeine
22	Acetylcodeine / 6-MAM / quinine	Acetylcodeine / 6-MAM / quinine	Acetylcodeine / 6-MAM
23	Quinine	Quinine	Galactose
24	Quinine	Quinine	Glucose
25	QUININE	QUININE	
26	Quinine	Quinine	
27	Quinine	Quinine	
28			
29	quinine, acetyl codeine, monoacetyl morphine	quinine, acetyl codeine, monoacetyl morphine	acetyl codeine, monoacetyl morphine, splenda
30	acetylcodeine, quinine	acetylcodeine, quinine	Glucose monohydrate
31	Quinine	Quinine	Glucose
32	quinine	quinine	
33			

*Some responses may be modified so that the participant cannot be identified.

6 DISCUSSION OF RESULTS

6.1 Assigned Value

The assigned value for the duplicate Samples S1 and S2 was determined as the robust average of the means of the results reported by each participant for these samples. The assigned value for Sample S3 was the robust averages of participants' results. The robust averages and associated expanded uncertainties were calculated using the procedure described in ISO 13528.[5] The calculation of the expanded uncertainty for a robust average is presented in Appendix 1.

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.

6.2 Measurement Uncertainty Reported by Participants

Participants were asked to report an evaluation of the expanded MU associated with their results, and the basis of this uncertainty evaluation. 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.[7]

Of 99 numeric results, 93 (94%) were reported with an associated expanded uncertainty. Participants used a wide variety of procedures to evaluate their reported uncertainties (Tables 3 and 4). One participant reported using the NATA GAG Estimating and Reporting MU as their guide; this document has been officially removed from the NATA website and is considered obsolete.[9]

Laboratories **6** and **25** did not report uncertainties for any of their results; both these participants reported that they were not accredited to ISO/IEC 17025.

The magnitudes of the reported uncertainties were within the range 2% to 45% relative to the reported result. In general, an expanded uncertainty of less than 3% relative may be unrealistically small for a routine measurement, while an expanded uncertainty of over 10% relative may be too large to be fit for purpose. Of the 99 expanded MUs reported, 3 were less than 3% relative, while 29 were greater than 10% relative.

Participants were also requested to report the coverage factor associated with their uncertainties (Table 3). Of the participants reporting coverage factors, most reported k = 2 at approximately 95% confidence level (16 participants). Five participants reported k = 3 at approximately 99% confidence level. One participant, Laboratory **17**, reported k = 1 at approximately 68% confidence level; therefore, this participant's uncertainties were reported as standard uncertainty rather than expanded uncertainty as requested for this PT study.

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

In some cases, results were reported with an inappropriate number of significant figures. Including too many significant figures may inaccurately reflect the precision of measurements. The recommended format is to write the 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 62.69 ± 3.13%, it is recommended to report 62.7 ± 3.1%.[8]

6.3 z-Score

The *z*-score compares the participants' deviation from the assigned value with the standard deviation for proficiency assessment (SDPA).

The SDPA defines acceptable performance in a proficiency test. Unlike the standard deviation based on between-laboratory CV, setting the SDPA to a realistic value enables *z*-scores to be used as fixed reference value points for assessment of laboratory performance, independent of group performance.

The between-laboratory CV predicted by the Thompson-Horwitz equation,[6] between-laboratory CV from reported results in this study, and the SDPA (as PCV) are presented for comparison in Table 11. SDPAs equivalent to 3% were used to calculate *z*-scores for all samples.

Table 11 Thompson-Horwitz CV, Between-Laboratory CV and SDPA (as PCV)

Sample	Analyte	Assigned Value (% base (m/m))	Thompson-Horwitz CV (%)	Between-Laboratory CV* (%)	SDPA (as PCV) (%)
S1	Heroin	61.3	1.3	4.5	3
S2	Heroin	61.3	1.3	4.5	3
S3	Heroin	27.2	1.9	5.6	3

*Robust between-laboratory CV with outliers removed, if applicable.

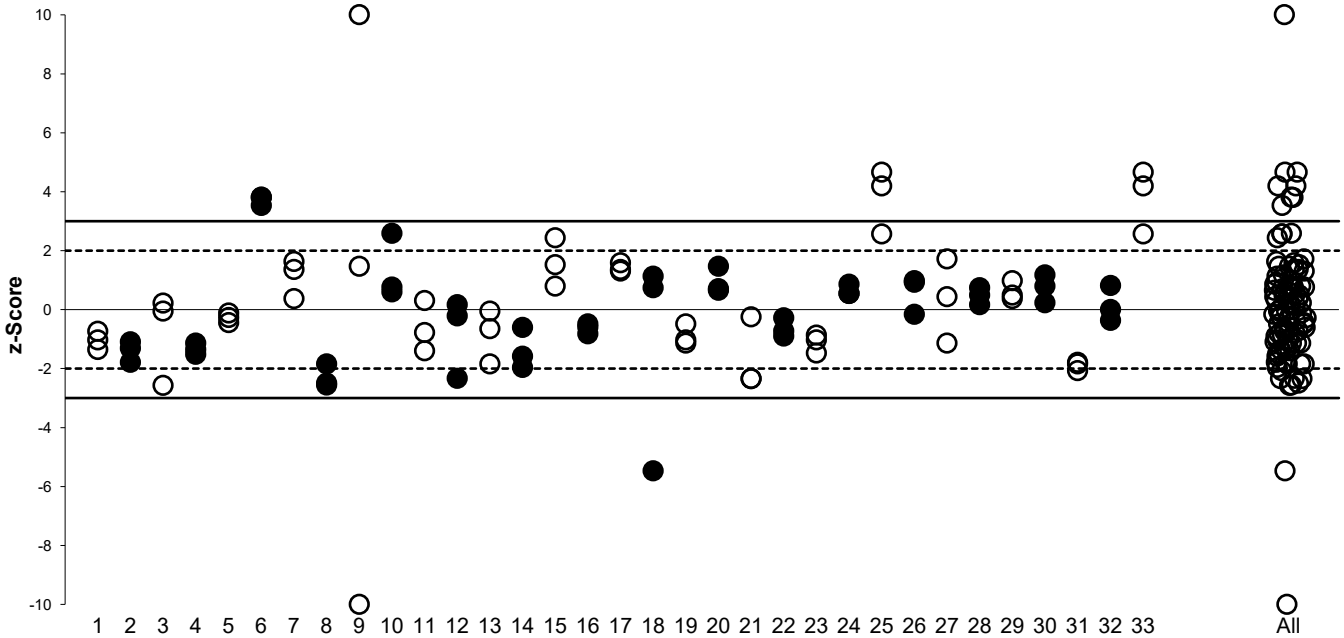
Of 99 results for which *z*-scores were calculated, 78 (79%) returned a *z*-score of $|z| \leq 2.0$, indicating an acceptable performance.

Laboratories **1**, **2**, **4**, **5**, **7**, **11**, **13**, **14**, **16**, **17**, **19**, **20**, **22**, **23**, **24**, **26**, **27**, **28**, **29**, **30**, and **32** returned acceptable z-scores for all reported numeric results.

Twelve participants returned at least one questionable or unacceptable *z*-score.

Laboratories **6**, **25** and **33** returned questionable or unacceptable *z*-scores for all samples (all positively biased). These participants may need to review their methodology to identify the source of this positive bias. Alternatively, these participants may have reported their results in units of % heroin salt (m/m) instead of % base (m/m).

The dispersal of participants' *z*-scores is presented graphically in Figure 5.



z-Scores greater than 10.0 or less than -10.0 have been plotted at 10.0 or -10.0 respectively.

Figure 5 *z*-Score Dispersal by Laboratory

6.4 E_n-Score

E_n-scores can be interpreted in conjunction with z-scores. The E_n-score indicates how closely a result agrees with the assigned value considering the respective uncertainties. An unacceptable E_n-score can be caused by inappropriate measurement, an inappropriate evaluation of measurement uncertainty, or both. Where a participant did not report an expanded uncertainty with a result, an expanded uncertainty of zero (0) was used to calculate the E_n-score.

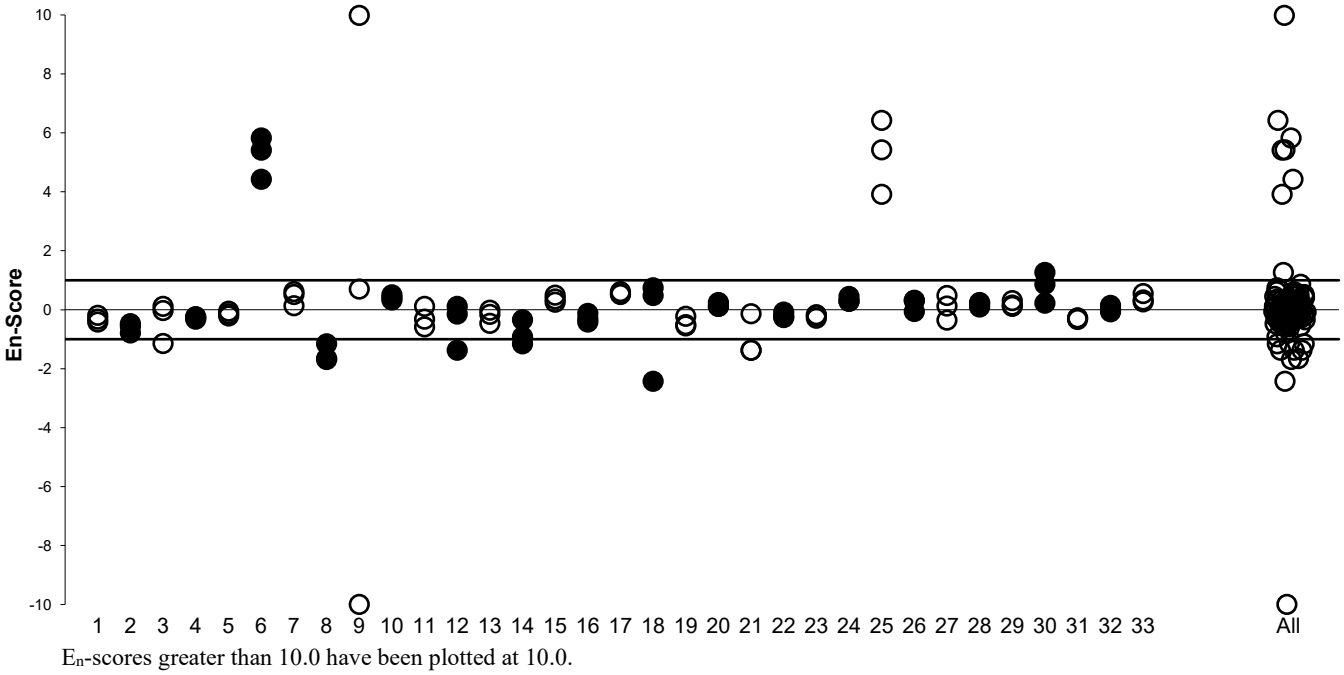
Of 99 results for which E_n-scores were calculated, 81 (82%) returned an E_n-score of $|E_n| < 1.0$, indicating agreement of the participant's result with the assigned value within their respective expanded uncertainties.

Laboratories **1**, **2**, **4**, **5**, **7**, **10**, **11**, **13**, **15**, **16**, **17**, **19**, **20**, **22**, **23**, **24**, **26**, **27**, **28**, **29**, **31**, **32**, and **33** returned acceptable E_n-scores for all reported numeric results.

Ten participants returned at least one unacceptable E_n-score.

Laboratories **6**, **8** and **25** returned unacceptable E_n-scores for all samples. Laboratories **6** and **25** did not report any uncertainties.

The dispersal of participants' E_n-scores is presented graphically in Figure 6.



E_n-scores greater than 10.0 have been plotted at 10.0.

Figure 6 E_n-Score Dispersal by Laboratory

6.5 Identification of Cutting Agents

Duplicate Samples S1 and S2 were prepared by adding quinine to heroin hydrochloride. Sample S3 was prepared by adding glucose to heroin hydrochloride.

Participants were requested to identify the cutting agents in the samples if part of their routine analysis, and the results reported are presented in Table 10.

Twenty-six participants (79%) reported on the identity of at least one cutting agent in the samples.

Six participants correctly identified all cutting agents in this study (Laboratories **1**, **8**, **11**, **24**, **30**, and **31**). Laboratory **3** correctly identified all cutting agents, however also reported an additional cutting agent for Sample S3 (lactose).

For duplicate Samples S1 and S2, 23 participants correctly identified quinine as the cutting agent (Laboratories **1**, **2**, **3**, **4**, **6**, **8**, **11**, **12**, **13**, **14**, **17**, **18**, **21**, **22**, **23**, **24**, **25**, **26**, **27**, **29**, **30**, **31**, and **32**). Laboratory **16** reported quinine in Sample S1, however did not report any cutting agents for Sample S2. Laboratory **9** reported quinine in Samples S1 and S3; this participant may have switched their results for Samples S2 and S3.

For Sample S3, eight participants correctly identified glucose as the cutting agent (Laboratories **1**, **3**, **8**, **10**, **11**, **24**, **30**, and **31**). Laboratory **3** also reported 'lactose' alongside glucose. Laboratories **4** and **14** reported general 'sugars' as the cutting agent. Laboratory **23** reported 'galactose', which is a stereoisomer of glucose. Laboratory **29** reported 'splenda', which contains sucralose, a compound structurally related to glucose. Laboratory **9** reported glucose for Sample S2 (as discussed above, this participant may have switched their results for Samples S2 and S3).

Across all three samples, some participants reported acetylcodeine and 6-monoacetylmorphine (6-MAM) as cutting agents. These are known impurities of heroin caused by degradation.

6.6 Participants' Analytical Methods

Participants were requested to analyse the samples using their routine test methods and to report a single result for each sample as they would normally report to a client. Results reported in this way reflect the true variability of results reported to laboratory clients. The method descriptions provided by participants are presented in Tables 1 and 2.

Plots of the *z*-scores versus various parameters are presented in Figures 7 to 10. A variety of sample masses, calibration standard sources, and extraction solvents were used by participants in this study. Results excluded from statistical calculations in Section 5 have also been excluded from discussion in this section.

Two laboratories reported using 50 – 100 mg of sample for analysis: Laboratories **8** and **33**. Laboratory **8** returned two questionable or unacceptable z-scores (both negatively biased). Laboratory **33** returned three questionable or unacceptable z-scores (all positively biased). These laboratories should consider reviewing their method bias. In general, 20 – 30 mg was the most commonly used sample mass by participants.

Instrumental techniques employed by participants for the analysis of heroin samples in this study included gas chromatography (GC) coupled with flame ionisation detection (FID) and liquid chromatography (LC) including high performance liquid chromatography (HPLC) or ultra performance liquid chromatography (UPLC) as reported by participants, coupled with diode array detection/photodiode array detection (DAD/PDA), UV-Vis, or tandem mass spectrometry (MS/MS). LC-DAD/PDA was the most common measurement instrument employed by participants. GC-FID was also a commonly used measurement instrument used by participants.

Laboratory **33** reported using LC-MS/MS for all samples and returned positively biased questionable or unacceptable *z*-scores for three of their reported results.

The measurement of heroin in Sample S3 posed a greater analytical challenge for participants than the measurement of heroin in duplicate Samples S1 and S2. This may be due to the lower level of heroin in this sample. The between-laboratory CV for duplicate Samples S1 and S2 was 4.5%, whilst for Sample S3 the between-laboratory CV was 5.6%.

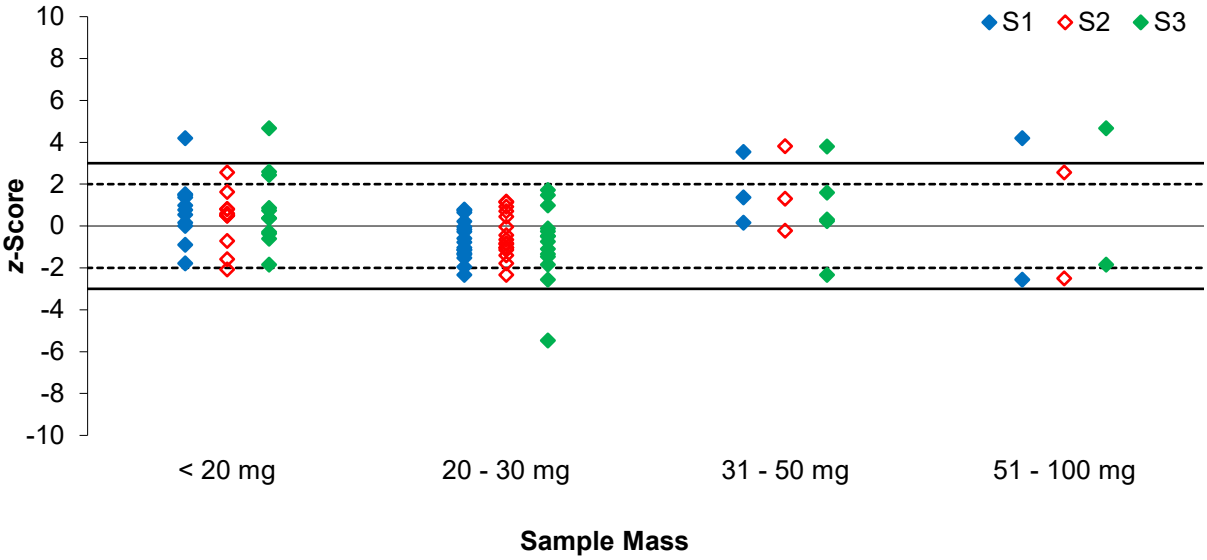


Figure 7 *z*-Score vs Sample Mass Used for Analysis

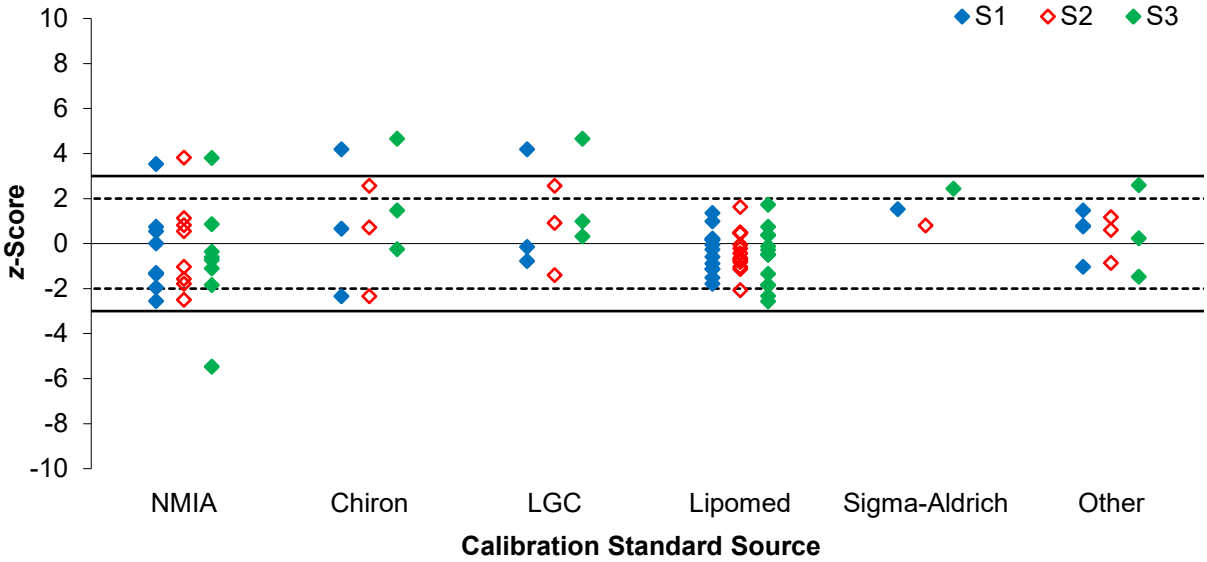


Figure 8 *z*-Score vs Calibration Standard Source

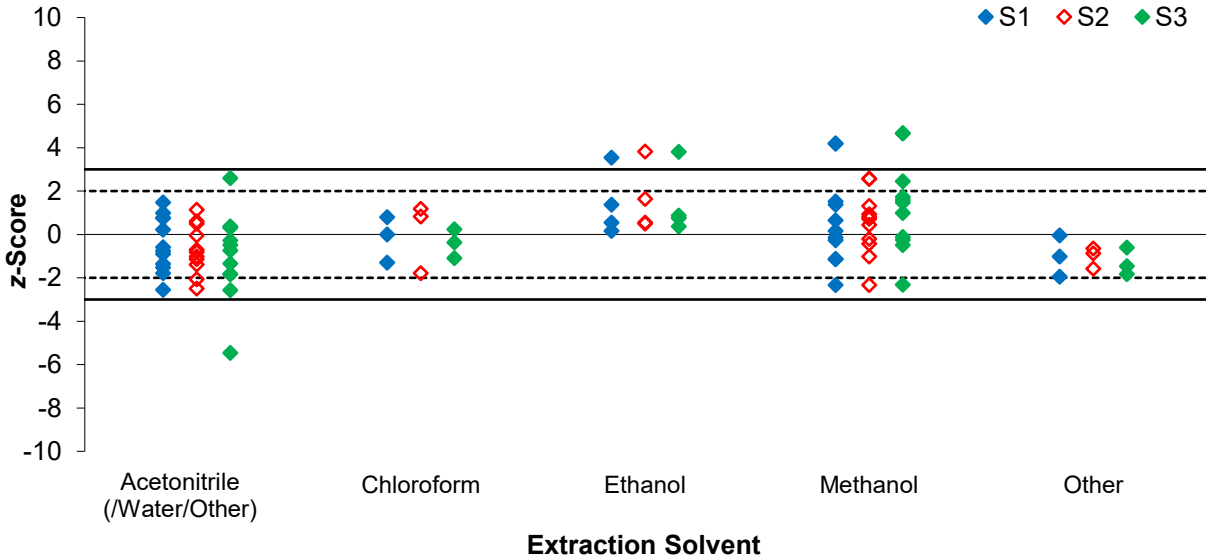


Figure 9 *z*-Score vs Extraction Solvent

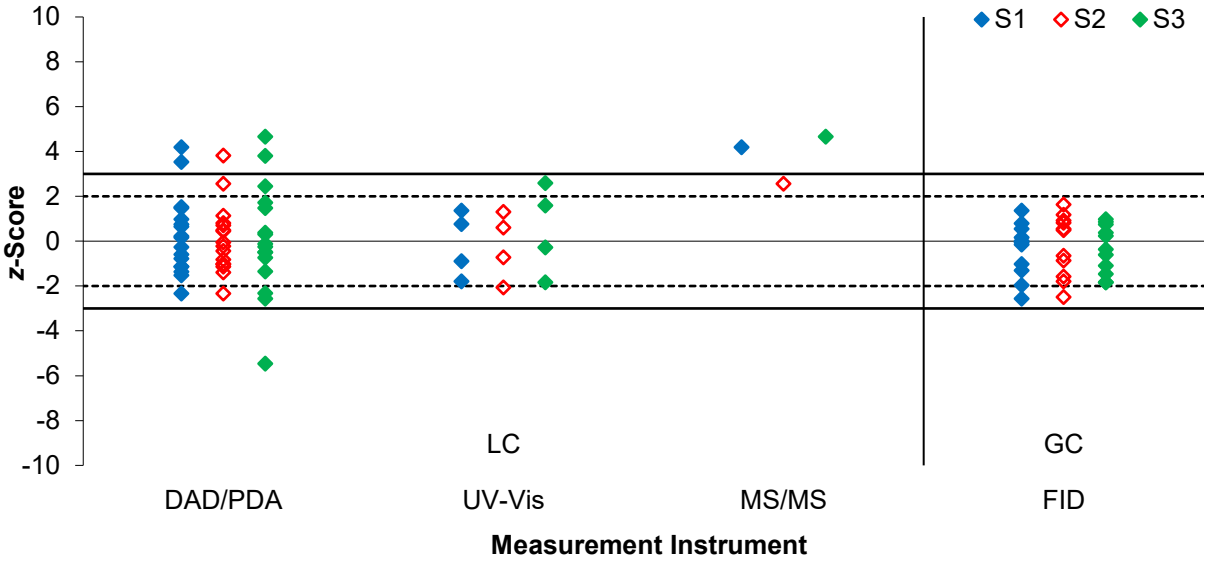


Figure 10 *z*-Score vs Measurement Instrument

6.7 Comparison of Results and Date of Analysis

As there were delays with sample delivery to some international participants, the samples were analysed by participants over approximately 4 months. No trend was found between when the samples were analysed and the results obtained (Figure 11).

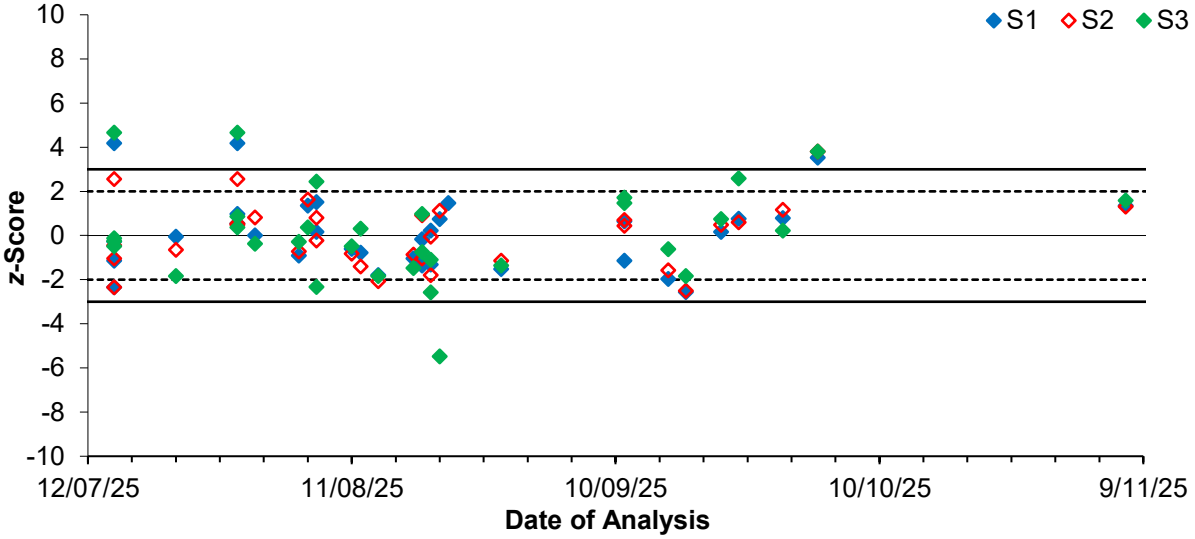


Figure 11 *z*-Score vs Sample Analysis Date

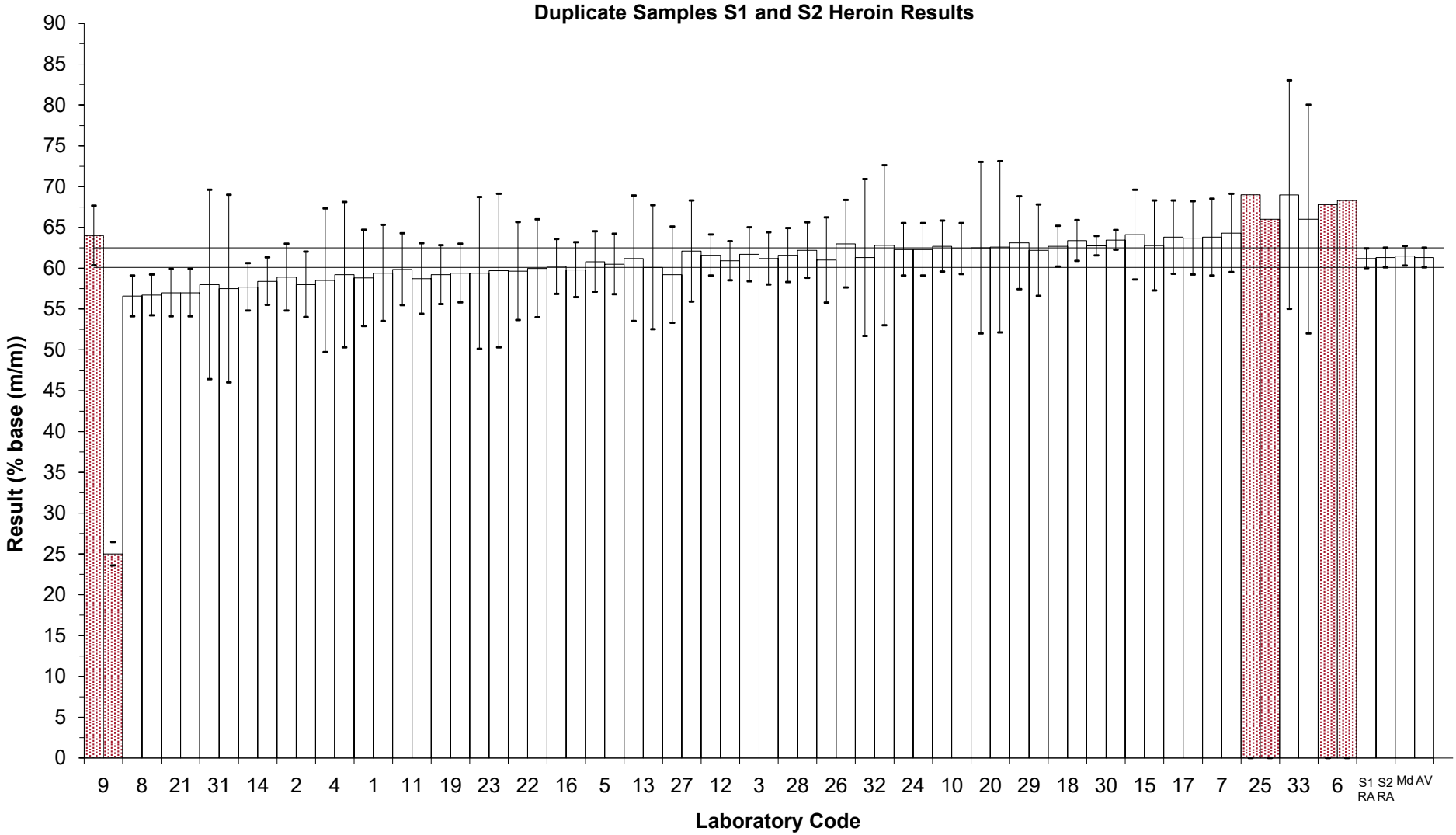
6.8 Duplicate Samples S1 and S2

Samples S1 and S2 were blind duplicate samples. Participants' results for these samples, along with the robust averages, median and assigned value, are presented in Figure 12.

The majority of participants' results for these samples were in agreement with each other within their reported expanded uncertainties, reflecting suitable within laboratory repeatability.

As Laboratories **6** and **25** did not report identical results, and also did not report uncertainties, their duplicate results were not in agreement. For both participants', their results may have been in agreement with each other if they had reported reasonable expanded uncertainties.

Based on the results and cutting agents reported by Laboratory **9** for Samples S2 and S3, it is suspected that this participant may have switched their Samples S2 and S3 results.



Horizontal lines are the assigned value ± U. Participants' results which are not in agreement with each other within reported uncertainties are shaded. RA = Robust Average, Md = Median, AV = Assigned Value.

Figure 12 Results for Blind Duplicate Samples S1 and S2

6.9 Comparison with Previous PT Studies

To enable direct comparison with previous NMIA heroin PT studies, the SDPA used to calculate *z*-scores has been kept constant at 3% PCV.

A summary of the acceptable performance, presented as a percentage of the total number of scores, obtained by PT study participants for Heroin from 2016 – 2025 (last ten studies) is presented in Figure 13. The average proportion of acceptable *z*-scores and E_n-scores over this period is 85% and 86% respectively. The samples in this study were challenging for participants compared to recent previous studies, with the lowest proportion of acceptable *z*-scores and E_n-scores since AQA 18-11.

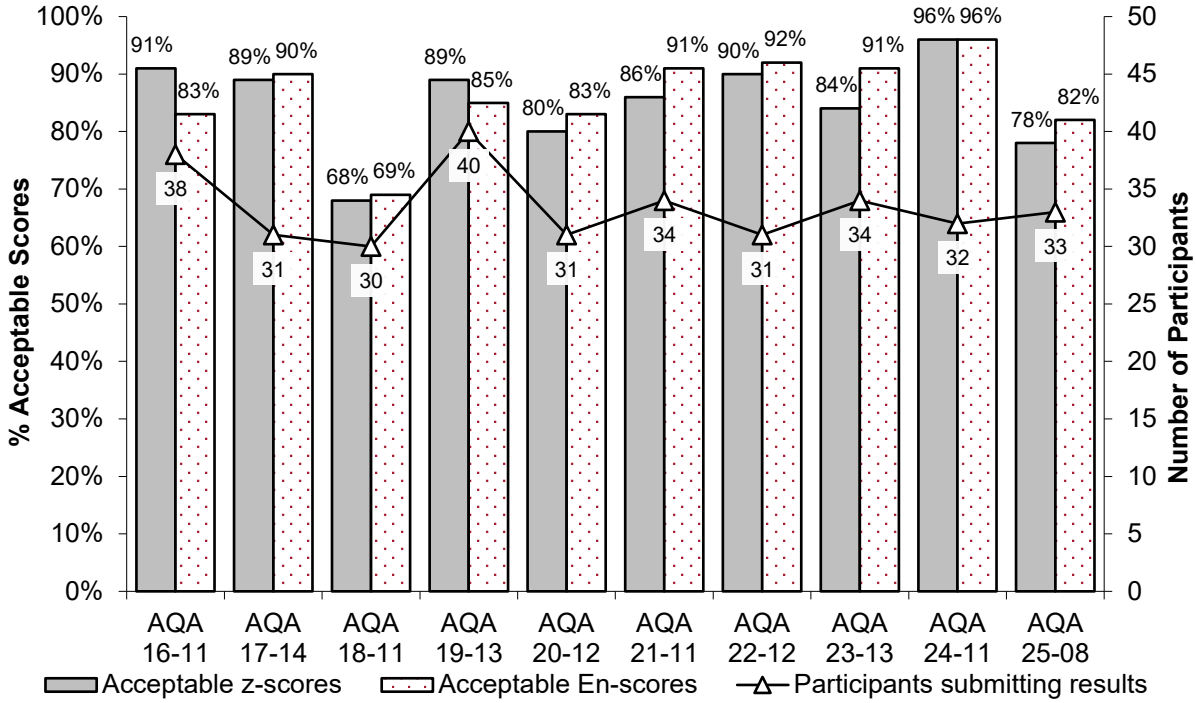


Figure 13 Summary of Participants' Performance in NMIA Heroin PT Studies

A number of participants have consistently participated in NMIA heroin PT studies, and individual performance history reports are emailed to each participant at the end of every PT study. The consideration of *z*-scores for an analyte over time provides much more useful information than a single *z*-score. Over time, laboratories should expect at least 95% of their scores to lie within the range $|z| \leq 2.0$. Scores in the range $2.0 < |z| < 3.0$ can occasionally occur, however, these should be interpreted in conjunction with the other scores obtained by that laboratory. For example, a trend of *z*-scores on one side of the zero line is an indication of method or laboratory bias.

A summary of individual laboratory's performances over the last ten NMIA heroin PT studies is presented in Figures 14 and 15 for Australian and international laboratories respectively. Two Australian and five international laboratories have achieved acceptable *z*-scores across all heroin samples in PT studies participated in over this period.

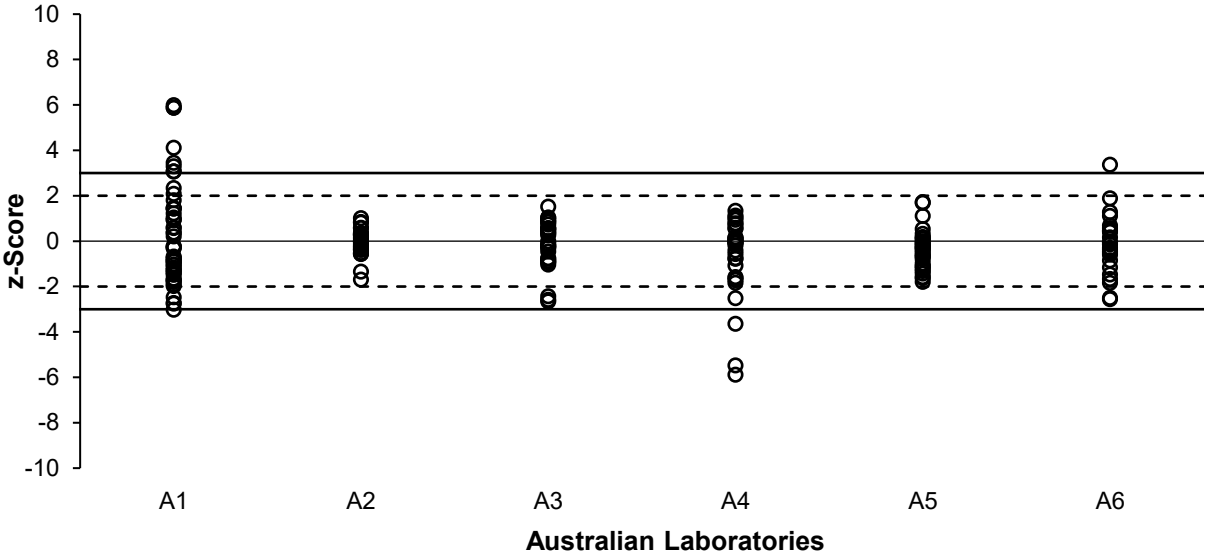
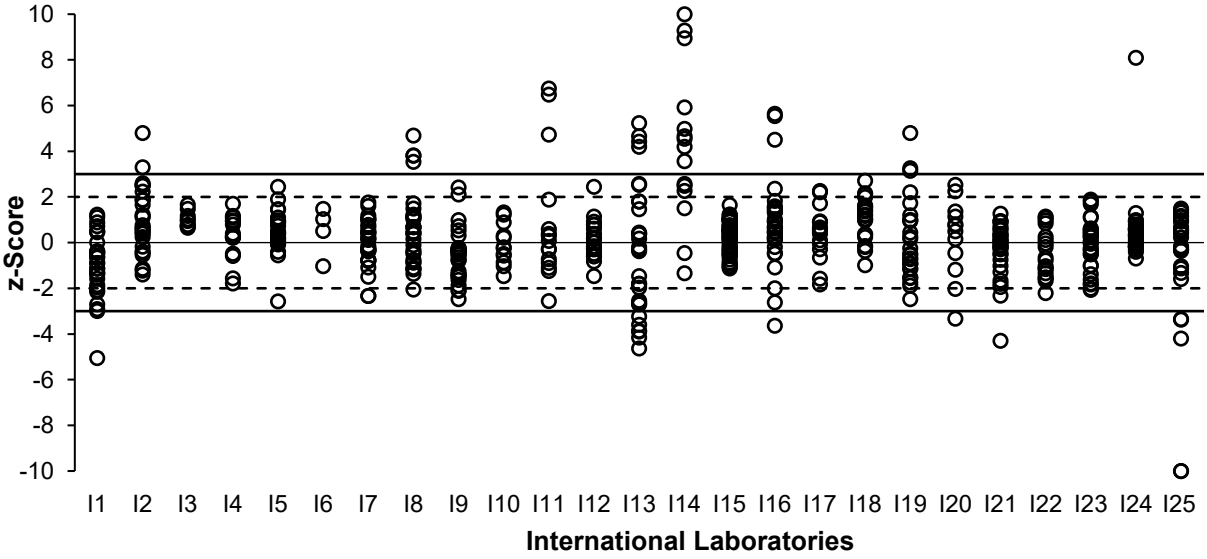


Figure 14 Summary of Australian Participants' *z*-Scores in NMIA Heroin PT Studies



z-Scores greater than 10.0 or less than -10.0 have been plotted at 10.0 or -10.0 respectively.

Figure 15 Summary of International Participants' *z*-Scores in NMIA Heroin 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, 2025, *Study Protocol for Proficiency Testing*, viewed November 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, 2025*, Chemical Proficiency Testing Statistical Manual*, viewed November 2025, <https://www.industry.gov.au/sites/default/files/2019-07/cpt_statistical_manual.pdf>.

[4] Thompson, M., Ellison, S.L.R. and Wood, R., 2006, 'The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories', *Pure Appl. Chem.,* vol. 78, pp. 145-196.

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

[6] 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.

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

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

[9] NATA, 2020, *Update to Measurement Uncertainty Resources*, viewed November 2025, <https://nata.com.au/news/update-to-measurement-uncertainty-resources/>

APPENDIX 1 ROBUST AVERAGE AND ASSOCIATED UNCERTAINTY, z-SCORE AND E_n-SCORE CALCULATIONS

A1.1 Robust Average and Associated Uncertainty

When the robust average is calculated using the procedure described in ISO 13528,[5] the uncertainty is evaluated as:

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

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 12.

Table 12 Uncertainty of Robust Average of Heroin in Sample S3

No. results (p)	32
Robust Average	27.2% base (m/m)
$S_{rob\ average}$	1.5% base (m/m)
$u_{rob\ average}$	0.33% base (m/m)
k	2
$U_{rob\ average}$	0.66% base (m/m)

Therefore, the robust average for Sample S3 heroin is 27.2 ± 0.7% base (m/m).

A1.2 z-Score and E_n-Score Calculations

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 13.

Table 13 z-Score and E_n-Score for Sample S3 Heroin Result Reported by Laboratory 20

Participant Result (% base (m/m))	Assigned Value (% base (m/m))	SDPA	z-Score	E_n-Score
28.4 ± 4.8	27.2 ± 0.7	3% as PCV, or: 0.03 × 27.2 = 0.816 % base (m/m)	$z = \frac{28.4 - 27.2}{0.816} = 1.47$	$E_n = \frac{28.4 - 27.2}{\sqrt{4.8^2 + 0.7^2}} = 0.25$

APPENDIX 2 ACRONYMS AND ABBREVIATIONS

CITAC	Cooperation on International Traceability in Analytical Chemistry
CRM	Certified Reference Material
CV	Coefficient of Variation
DAD	Diode Array Detection
EA	European Accreditation
FID	Flame Ionisation Detection
GAG	General Accreditation Guidance (NATA)
GC	Gas Chromatography
GUM	Guide to the expression of Uncertainty in Measurement
HPLC	High Performance Liquid Chromatography
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
k	Coverage factor
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
NMIA	National Measurement Institute Australia
NR	Not Reported
PCV	Performance Coefficient of Variation
PDA	Photodiode Array
PT	Proficiency Testing
RA	Robust Average
RM	Reference Material
SD	Standard Deviation
SDPA	Standard Deviation for Proficiency Assessment
SI	International System of Units
SS	Spiked Samples
UPLC	Ultra Performance Liquid Chromatography
UV/Vis	Ultraviolet/Visible spectroscopy

END OF REPORT