JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, January 2010, Vol. 55, No. 1
doi: 10.1111/j.1556-4029.2009.01242.x
Available online at: interscience.wiley.com

TECHNICAL NOTE

CRIMINALISTICS

Sarah J. Benson,[1,2] *Ph.D.; Christopher J. Lennard,*[1,3] *Ph.D.; Philip Maynard,*[2] *Ph.D.; David M. Hill,*[4] *B.Sc.; Anita S. Andrew,*[5] *Ph.D.; Ken Neal,*[6] *B.Sc.; Hilary Stuart-Williams,*[7] *Ph.D.; Janet Hope,*[8] *B.Sc.; G. Stewart Walker,*[9] *Ph.D.; and Claude Roux,*[2] *Ph.D.*

Forensic Analysis of Explosives Using Isotope Ratio Mass Spectrometry (IRMS)—Part 2: Forensic Inter-Laboratory Trial: Bulk Carbon and Nitrogen Stable Isotopes in a Range of Chemical Compounds (Australia and New Zealand)

ABSTRACT: Comparability of data over time and between laboratories is a key issue for consideration in the development of global databases, and more broadly for quality assurance in general. One mechanism that can be utilized for evaluating traceability is an inter-laboratory trial. This paper addresses an inter-laboratory trial conducted across a number of Australian and New Zealand isotope ratio mass spectrometry (IRMS) laboratories. The main objective of this trial was to determine whether IRMS laboratories in these countries would record comparable values for the distributed samples. Four carbon containing and four nitrogen containing compounds were distributed to seven laboratories in Australia and one in New Zealand. The laboratories were requested to analyze the samples using their standard procedures. The data from each laboratory was evaluated collectively using International Standard ISO 13528 (*Statistical methods for use in proficiency testing by inter-laboratory comparisons*). "Warning signals" were raised against one participant in this trial. "Action signals" requiring corrective action were raised against four participants. These participants reviewed the data and possible sources for the discrepancies. This inter-laboratory trial was successful in providing an initial snapshot of the potential for traceability between the participating laboratories. The statistical methods described in this article could be used as a model for others needing to evaluate stable isotope results derived from multiple laboratories, e.g., inter-laboratory trials/proficiency testing. Ongoing trials will be conducted to improve traceability across the Australian and New Zealand IRMS community.

KEYWORDS: forensic science, isotope ratio mass spectrometry, inter-laboratory trial, traceability, nitrogen isotope, carbon isotope, isotope ratios

The potential for the isotope ratio mass spectrometry (IRMS) technique to be utilized in the forensic field to assist in identifying the source of a variety of evidence types (e.g., explosives and illicit drugs) is well recognized and has recently been reviewed (1). Given the relatively short history of use of IRMS in the forensic science industry, the technique requires validation for a range of applications prior to casework implementation. One aspect of the technique requiring validation is traceability—the comparability of results obtained by different laboratories and at different times (2). In order to evaluate traceability in the Australian and New Zealand IRMS community, an inter-laboratory trial was organized. Samples were prepared and distributed to a number of Australian and New Zealand IRMS laboratories. A number of international inter-laboratory trials have been conducted, including those by the FIRMS Network (3–5); however, it was deemed necessary to conduct a locally based trial as the laboratories in the region would be the ones most likely to exchange data and share databases.

The primary objective of this inter-laboratory trial was to determine whether IRMS laboratories in Australia and New Zealand would report comparable values for the distributed samples. A secondary objective was to assist in determining the true stable isotope values for laboratory standards for use in the Australian Federal Police laboratory for the determination of bulk nitrogen and carbon isotope ratios. Additionally, the compilation of details of methodology used in each laboratory would provide data to recommend best

[1]Forensic & Data Centres, Australian Federal Police, GPO Box 401, Canberra 2601, ACT, Australia.
[2]Centre for Forensic Science, University of Technology, Sydney, PO Box 123, Broadway 2007, NSW, Australia.
[3]National Centre for Forensic Studies, University of Canberra, Canberra 2601, ACT, Australia.
[4]Australian Nuclear Science and Technology Organisation, PMB 1, Menai 2234, NSW, Australia.
[5]Environmental Isotopes Pty. Ltd., PO Box 1492, Macquarie Centre 2113, NSW, Australia.
[6]Iso-trace Ltd., Centre for Innovation, St David Street, Dunedin, New Zealand.
[7]Research School of Biological Sciences, Australian National University, GPO Box 475, Canberra 2601, ACT, Australia.
[8]Geoscience Australia, GPO Box 378, Canberra 2601, ACT, Australia.
[9]School of Chemistry, Physics and Earth Sciences, Flinders University, GPO 2100, Adelaide 5001, SA, Australia.

Received 6 Oct. 2007; and in revised form 17 May 2008; accepted 24 Jan. 2009.

practice with respect to standard analytical sequences, international and laboratory standards, methods, and correction calculations for this specific application of IRMS.

Materials and Methods

Sample Preparation and Distribution

The following samples were distributed to seven laboratories in Australia and one laboratory in New Zealand, with a request to analyze the samples using the laboratory's standard procedures and methods for bulk carbon isotope values ($\delta^{13}C_{VPDB}$ (‰)) and bulk nitrogen isotope values ($\delta^{15}N_{AIR}$ (‰)).

- Carbon samples: calcium carbonate, caffeine, sucrose, and plain flour.
- Nitrogen samples: ammonium nitrate, potassium nitrate, ammonium thiocyanate, and plain flour.

The eight selected laboratories are representative of the IRMS laboratories in Australia and New Zealand.

The samples were selected following a preliminary evaluation of potentially suitable materials to serve as laboratory standards for the measurement of bulk nitrogen and carbon isotope values in inorganic samples, particularly ammonium nitrate, and a range of samples containing carbon. The evaluation focused on: similarity in chemical composition, whether the isotope values of the potential materials bracketed the expected isotope values of the target materials, homogeneity, availability, similarity of decomposition products in pyrolysis cycles, and ease of use.

The following information was also requested from the participating laboratories: details of instrumentation (i.e., manufacturer and model); details of standards including their true/certified and measured values during the trial (international and/or laboratory standards); weight of each sample analyzed; correction calculations performed on the data produced from the instrument; quality control checks performed to ensure reliability of data; copies of sequences; and a summary of methods for carbon and nitrogen bulk measurements.

The samples were prepared by grinding a small quantity of each sample to a fine consistency. No attempt was made to size the particles; it is intended that this will be conducted in the next round. The flour and calcium carbonate samples were not ground as these already had a fine consistency. The mortar and pestle were prewashed with deionized water and acetone, oven dried at 50°C for 10 min, and allowed to cool. A small sample of each solid was then transferred to a 4 mL glass vial fitted with a Teflon-lined screw cap (Sigma-Aldrich, St. Louis, MO). The screw caps were sealed with Parafilm. The samples, secured in a box with a desiccator cartridge (Sigma-Aldrich), were then distributed to participating laboratories. The trial guidelines and relevant material safety data sheets were also enclosed. A timeframe of *c.* 6 weeks was specified for completion of the trial (however some results were not submitted for 12 months).

Application of International Standard ISO 13528:2005(E) Statistical Methods for Use in Proficiency Testing by Inter-laboratory Comparisons

The trial results were compiled and evaluated in accordance with International Standard ISO 13528:2005(E) *Statistical methods for use in proficiency testing by inter-laboratory comparisons* (6) and subsequently distributed to participants for comment.

TABLE 1—*Summary of the mean ($\delta^{13}C_{VPDB}$ (‰)) and standard deviation (s) data for the samples distributed for carbon measurement. The number of samples (No.) measured by each participant is also provided. Nil response from laboratory 2.*

	Calcium Carbonate			Caffeine			Sucrose			Plain Flour		
Lab.	Mean	*s*	No.	Mean	*s*	No.	Mean	*s*	No.	Mean	*s*	No.
1	−43.51	0.36	10	−40.38	0.04	10	−11.28	0.03	10	−22.39	0.04	10
3	−41.34	0.16	10	−39.11	0.21	10	−9.58	0.37	10	−21.68	0.72	10
4	−42.87	0.20	6	−40.93	0.17	6	−11.27	0.02	4	−22.65	0.11	4
5	−42.44	0.28	7	−40.87	0.06	7	−11.64	0.24	5	−22.56	0.06	7
6	NR*	NR*	NR*	−39.79	0.86	4	−11.39	0.02	2	−22.44	0.07	4
7	−42.30	0.76	4	−40.73	0.11	4	−11.51	0.15	3	−22.74	0.22	4
8	NR*	NR*	NR*	−40.22	0.03	10	−11.82	0.10	10	−22.46	0.09	10
Mean		−42.49			−40.29			−11.21			−22.42	
s		0.80			0.66			0.74			0.35	

*Indicates that no result/data was reported.

TABLE 2—*Summary of the mean ($\delta^{15}N_{AIR}$ (‰)) and standard deviation (s) data for the samples distributed for nitrogen measurement. The number of samples (No.) measured by each participant is also provided. Nil response from laboratory 2.*

	Ammonium Nitrate			Potassium Nitrate			Ammonium Thiocyanate			Plain Flour		
Lab.	Mean	*s*	No.	Mean	*s*	No.	Mean	*s*	No.	Mean	*s*	No.
1	0.72	0.13	10	−0.64	0.50	10	3.48	0.07	10	5.42	0.17	10
3	−0.14	0.38	10	−0.41	0.40	10	2.76	0.18	10	4.49	0.80	5
4	−0.17	0.05	3	−0.44	0.05	4	2.83	0.06	3	5.78	0.15	3
5	−0.27	0.07	7	−0.35	0.03	7	2.93	0.06	7	3.65	0.10	6
6	0.13	0.01	2	−0.06	0.19	2	3.15	0.05	2	NR*	NR*	NR*
7	−1.39	0.16	3	−0.70	0.39	1	2.94	0.04	4	4.42	0.94	4
8	−0.01	0.05	10	−0.41	0.15	10	3.29	0.21	10	4.82	0.37	8
Mean		−0.16			−0.43			3.05			4.76	
s		0.63			0.21			0.26			0.76	

*Indicates that no result/data was reported.

TABLE 3—*Summary of the analytical methods and instrument configurations utilized by each participant. Nil response from laboratory 2.*

Lab.	Instrument	Reference Gas	Carrier Gas	Gas Storage	Oven Temp (°C)	Combustion Reactor Temp (°C)	Combustion Reactor Packing	Reduction Reactor Temp (°C)	Reduction Reactor Packing	Water Trap
1	Sercon 20/20 differentially pumped IRMS coupled to a Carlo Erba NA1500	CO_2 and N_2	He UHP	Air conditioned lab	60	1000	Chromium oxide, copper oxide wire, silver wire	600	Cu	$MgClO_4$
3	GV Instruments IsoPrime EuroVector EA	CO_2 UHP, N_2 UHP	He UHP	Air conditioned lab	NR*	1030	Silvered cobaltous/cobaltic oxide, copper oxide also CO_2 trap used for N_2 analysis	650	Reduced Cu	$MgClO_4$
4	MAT252 (Thermo Finnigan) IRMS, with ConFlo III interface and modified Roboprep EA	N_2 UHP (99.999%), O_2 (99.996%), CO_2 (Food Grade)	He UHP	Air conditioned lab (approx 20°C)	50	1020	Chromium oxide, copper oxide wire, silver wire	550	H-reduced Cu	$MgClO_4$
5	DELTA^{plus}XP (Thermo Fisher) IRMS, ConFlo III FlashEA^{TM} 1112	N_2 UHP, O_2 (99.996%), CO_2 UHP	He UHP	Outside (not temp. controlled)	35	900	Silvered cobaltous/cobaltic oxide, copper oxide	680	Reduced Cu	$MgClO_4$
6	Thermo FlashEA, Conflo III, MAT 252	N_2, CO_2	He UHP	Air conditioned lab	40	900	Copper oxide/silvered cobaltous oxide	650	Reduced Cu	$MgClO_4$
7	CE EA1110 Micromass Isochrom+diluter	N_2, O_2 and CO_2 UHP	He UHP	Outside (not temp. controlled)	22	1020	Chromium oxide, silvered cobaltous oxide	660	Cu	$MgClO_4$
8	GV Instruments IsoPrime EuroVector EA	CO_2 UHP, N_2 UHP	He UHP	Outside (not temp. controlled)	40	1030	Chromium oxide, silvered cobaltous oxide	650	Reduced Cu	$MgClO_4$

*Indicates that no result/data was reported.

Determination of the Assigned Value

The assigned value (X) is the value attributed to a sample accepted as having an uncertainty appropriate for a given purpose. The assigned value for a material can be determined utilizing a number of methods and will depend on the number of participants, what is being determined in the trial and whether the test material is: prepared by mixing proportions of constituents; a certified reference material; or, measured against certified reference materials by one laboratory. In this trial, the assigned value was determined by utilizing the consensus value from the participants. This was utilized rather than other methods as the samples were not certified reference materials or mixtures; the assigned value was not previously determined through calibration with certified reference materials or through consensus values from expert laboratories.

Determining the assigned value using the consensus value from the participants means it is the robust average of the results reported by all participants in the round. For this application, robust average refers to an estimate of the population mean calculated using a robust algorithm. Refer to the Appendix for an outline of the algorithm used to calculate the assigned value.

Determination of the Standard Deviation for Proficiency Assessment

The standard deviation for proficiency assessment (σ) is utilized to assess the size of estimates of laboratory bias found in a trial. In this trial, the standard deviation was calculated using the participant's data and is the robust standard deviation of the results reported by the participants (i.e., estimate of the population standard deviation calculated using a robust algorithm) (6). Refer to the Appendix for an outline of the algorithm used to calculate the standard deviation for proficiency assessment.

Determination of the Standard Uncertainty of the Assigned Value

The standard uncertainty of the assigned value (μ_X) depends on the method employed to derive it, the number of participants, and other factors (6). In this trial, the assigned value was derived as a robust average, and as a result the uncertainty of the assigned value is estimated using the formula in the Appendix (step 8).

If the standard uncertainty of the assigned value is too large in comparison with the standard deviation for proficiency testing, then there is a risk that some laboratories will receive action and warning signals due to the inaccuracy of the assigned value, rather than due to an issue within the laboratory. According to ISO 13528, if the standard uncertainty of the assigned value (μ_X) is $>0.3\sigma$, then the uncertainty is significant and should be included in the interpretation of the results of the proficiency test. If the uncertainty is significant, the trial organizer must consider:

- Alternative methods for determining the assigned value where the uncertainty may meet the aforementioned criteria;
- Utilizing the uncertainty of the assigned value in the interpretation of the results; and
- Informing the participants that the uncertainty of the assigned value is not negligible.

Calculation of Performance Statistics

The participant's results were transformed into performance statistics to assist with the interpretation and comparison of trial data.

TABLE 4—*Details regarding sample storage, format of the analytical sequences utilized by each participant, and brief details regarding correction calculations performed on the data. Nil response from laboratory 2.*

Lab.	Sample Storage	Standard Sequence Format	Correction Calculations
1	Samples stored in the original vials	2 of each Std. followed by 10 or 12 samples depending on batch size and then an internal reference (e.g., EDTA) with a set of international stds. at end	All data corrected to IAEA stds
3	Glass desiccator silica gel. Lab. stds.—glass/Teflon. IAEA stds.—original vials	No result/data was reported	No result/data was reported
4	Stainless steel desiccator box with self-indicating silica gel. Lab. stds.—glass vials. IAEA stds.—original vials	1 × blank, 2 × Std. C1, 2 × Std. C2, 2 × Std. N1, 2 × Std. N2, 1 × concentration calibration, 12 × sample, repeat 1 × Stds.	1. Drift correction if required. 2. Slope and intercept correction using standards
5	Perspex & glass desiccators with self-indicating orange silica gel. Lab. stds.—glass vials with Teflon-lined screw caps. IAEA stds.—original vials	2 × blanks, 3 × Std. 1, 3 × Std. 2, 18 × Samples, 3 × Std. 1, 3 × Std. 2	1. Outliers excluded using Grubbs. 2. Calculate required correction (slope and intercept) by plotting true versus measured standard values. 3. Calculate true value of samples and standards utilizing equation: slope × measured value + intercept
6	In original vials	2 × blanks, 2 × Std. 1, 1 × Std. 2, 2 × Sample 1, 1 × Std. 3, 1 × blank, 2 × Sample 2, 1 × Std. 1	Carbon—corrected using TO_2 std. Normalized using TO_2 and caffeine stds. Nitrogen—corrected using caffeine std. Normalized using caffeine and IAEA stds
7	In original vials	Typically 3 × Std. 1, 3 × Std. 2, ~20 samples, repeat Stds. × 2	No statistical outlier removal—outliers removed very reluctantly by size or sizeable combined isotope anomaly relative to other samples, after setting cutoffs for whole run. Slope and intercept correction with multiple stds
8	In original vials, N samples in desiccator	Typically 2 × blanks, 2 × Std., 8 × samples, 1 × Std., repeat sequence	1. Drift correction for Nitrogen. 2. Slope and intercept correction using standards

The performance statistics for the trial were calculated using two methods—estimate of laboratory bias and *z*-scores. The estimate of laboratory bias is the difference between the participant's result and the assigned value. The *z*-scores provide a standardized measure of laboratory bias and are calculated using the assigned value and the standard deviation for proficiency assessment. Refer to the Appendix for the formulae utilized (steps 9 and 10).

The performance statistics for the trial were summarized graphically in the form of bar-plots of the standardized laboratory bias (*z*-scores).

Evaluation of the Data—Action and Warning Signals

According to ISO 13528, action signals are awarded when results are so incomparable that they merit investigation and corrective action. One action signal in one round of a trial or two warning signals in successive rounds are considered evidence that an anomaly has occurred that requires investigation (6).

If the laboratory bias is $>3.0\sigma$ or $<-3.0\sigma$ then the participant is awarded an action signal for that result. If the laboratory bias is $>2.0\sigma$ or $<-2.0\sigma$ then the participant is given a warning signal (6).

Z-scores above 3.0 or below −3.0 are awarded action signals and *z*-scores above 2.0 or below −2.0 are awarded warning signals (6).

Results and Discussion

A summary of the means and standard deviations of the carbon and nitrogen isotope values for the distributed samples are displayed in Tables 1 and 2. Table 3 provides a summary of the methods and instrument configurations utilized by each participant, and Table 4 summarizes the details regarding sample storage, format of the analytical sequences utilized by each participant, and brief details regarding correction calculations performed on the data. Table 5 provides details regarding the certified and measured values of analytical standards. Of particular interest are the reported certified values for the carbon international standards, some of which correspond with the recently updated consensus values as reported by Coplen et al. 2006 (7). Participants were not requested to measure the carbon isotope in ammonium thiocyanate or the nitrogen isotope in caffeine. A number of laboratories did measure and report similar values for these isotopes; however, as the majority of laboratories did not report results, these values were not evaluated in this study.

The calculated assigned value (X) for the carbon samples, standard uncertainty of assigned value (μ_X), standard deviation for proficiency assessment (σ), and whether the uncertainty of the assigned value was negligible for the carbon samples are displayed in Table 6. Table 7 displays the corresponding values for the nitrogen samples.

As $\mu_X > 0.3\sigma$ for the carbon and nitrogen samples, the participants were informed that the uncertainty of the assigned value was significant. The participants will investigate potential causes within their own laboratories which may have contributed to the significant uncertainty. This value will be compared against the uncertainty of the assigned value in future trials.

Refer to Table 8 for a summary of the performance statistics (estimates of laboratory bias and *z*-scores respectively) for the round for each sample.

The following action and warning signals were highlighted during interpretation of the laboratory bias and *z*-score results in Table 8.

- Laboratory 1—action signal for the nitrogen isotope value of ammonium nitrate;

TABLE 5— *Details regarding analytical standards utilized during the trial (both international and laboratory standards). Nil response from laboratory 2.*

	Nitrogen Standards					Carbon Standards				
	International Standards		Laboratory Standards			International Standards		Laboratory Standards		
Lab.	Name	$\delta^{15}N_{AIR}$ (‰)	Name	$\delta^{15}N_{AIR}$ (‰)	Weight (μg)	Name	$\delta^{13}C_{VPDB}$ (‰)	Name	$\delta^{13}C_{VPDB}$ (‰)	Weight (μg)
1	IAEA-N1, IAEA-N2	No result	EDTA	−0.9	100	IAEA-CH-7, IAEA-CH-6	No result	EDTA	−38.3	400
3	IAEA-N1, IAEA-N2	No result	Caffeine	−5.1	450	IAEA-CH-6, NBS 22	No result	Sugar	−11.9	120
4	IAEA-N1, IAEA-N2, IAEA-N3, USGS25, USGS26	0.43, 20.41, 4.72, −30.4, 53.7	B2036 B2105 Atropine	−3.42 ± 0.17 (n = 23) −5.08 ± 0.13 (n = 137) −14.37 ± 0.19 (n = 61)	1000 1000 1000	NBS 19, NBS 22, USGS24, IAEA-CH-7, IAEA-CH-6, L-SVEC	1.95, −30.03, −16.05, −32.15, −10.3, −46.6	B2036 B2105 Atropine	−27.41 ± 0.15 (n = 28) −28.43 ± 0.18 (n = 125) −30.38 ± 0.17 (n = 58)	1000 1000 1000
5	IAEA-N1, IAEA-N2, USGS25, IAEA-N3	0.4, 20.3, −30.4, 4.7	Ammonium nitrate Ammonium thiocyanate	−2.90 ± 0.30 2.90 ± 0.25	300 260	IAEA-CH-7, IAEA-CH-6, L-SVEC, NBS 19	−31.8, −10.4, −46.5, 1.95	Sucrose Ammonium thiocyanate	−11.48 ± 0.07 −31.88 ± 0.04	100 260
6	IAEA-N1, IAEA-N2	0.43, 20.41	Caffeine	−9.62	300	IAEA-CH-6	No result	Caffeine, TO2	−36.88, −22.98	100, 50
7	Oxford alanine, USGS40, USGS41, NIST 1547, IAEA-600, IAEA N3	−1.57, −4.50, 47.60, 2.02, 1.00, 4.70	Sea-line (nylon line) Gelatine Glucosamine	2.63 ± 0.29 8.35 ± 0.18 −1.69 ± 0.18	1000 ± 250	IAEA-CH-6, Oxford alanine, USGS40, USGS41, NIST 1547, IAEA-600	−10.45, −27.00, −26.39, 37.80, −25.89, −27.77	Beet Sea-line	−24.62 ± 0.05 −28.22 ± 0.15	1000 ± 250
8	IAEA-N1, USGS25, IAEA-N3	0.4, −30.4, 4.7	B2157	2.85	400	IAEA-C8 (internal), IAEA-CH-5	−18.31, −10.4			200

TABLE 6—*Calculated assigned value, standard uncertainty of assigned value, standard deviation for proficiency assessment, and whether the uncertainty of the assigned value is negligible for distributed carbon samples.*

	Calcium Carbonate	Caffeine	Sucrose	Plain Flour
Robust average/assigned value, X (‰)	−42.5	−40.3	−11.4	−22.5
Robust std deviation/std deviation for proficiency assessment, σ (‰)	0.8	0.7	0.2	0.2
Standard uncertainty of assigned value, μ_X (‰)	0.5	0.3	0.1	0.1
Is $\mu_X \leq 0.3\sigma$	No	No	No	No

TABLE 7—*Calculated assigned value, standard uncertainty of assigned value, standard deviation for proficiency assessment, and whether the uncertainty of the assigned value is negligible for distributed nitrogen samples.*

	Ammonium Nitrate	Potassium Nitrate	Ammonium Thiocyanate	Plain Flour
Robust average/assigned value, X (‰)	−0.1	−0.4	3.0	4.8
Robust std deviation/std deviation for proficiency assessment, σ (‰)	0.2	0.1	0.3	1.0
Standard uncertainty of assigned value, μ_X (‰)	0.1	0.1	0.1	0.5
Is $\mu_X \leq 0.3\sigma$	No	No	No	No

- Laboratory 3—action signal for the carbon isotope values of sucrose and plain flour;
- Laboratory 6—action signal for the nitrogen isotope value of potassium nitrate;
- Laboratory 7—action signal for the nitrogen isotope value of ammonium nitrate and warning signal for the nitrogen isotope value of potassium nitrate.

As this was the first round of the trial, the single warning signal cannot be regarded as an anomaly. The laboratories issued with action signals were provided with the opportunity to review their responses and provide possible reasons for the discrepancies. Corrective actions for consideration include: checking that users follow standard procedures; check that the standard procedures are correct; check the calibration of the instrument and composition of standards/samples; and/or comparative tests of users, instruments, and samples with another laboratory. The effectiveness of proposed corrective actions will be evaluated in future trials.

The performance statistics for the round were summarized graphically in the form of bar-plots of the standardized laboratory bias (z-scores) (Fig. 1*a*,*b*). The bar plots assist in identifying results that may be worthy of further investigation, e.g., are any participants producing large positive z-scores for all measurements (possible systematic errors) (6). Upon examination of the bar plots, no participants were producing large positive or large negative z-scores for all measurements; however, this will be monitored in future trials.

Conclusions

This inter-laboratory trial was successful in providing an initial snapshot of the potential for traceability between seven Australian and New Zealand IRMS laboratories with respect to reported carbon and nitrogen bulk stable isotope values. The statistical methods described in this article could be used as a model by other practitioners evaluating stable isotope results derived from multiple laboratories e.g., inter-laboratory trials. The trial was valuable in collating procedural details that will assist in determining best practice for stable isotope ratio measurement in the forensic science community. The trial also assisted with the determination of the true stable isotope values for the samples that were selected as laboratory standards for use in the Australian Federal Police laboratory (refer to Table 9).

Overall, a ''warning signal'' was raised against one participant. ''Action signals'' requiring corrective action were raised against four participants. Laboratories reviewed their data and considered potential corrective actions. Ongoing trials will be conducted to improve traceability across the Australian and New Zealand IRMS community.

TABLE 8—*Performance statistics: estimates of (a) laboratory bias and (b) z-scores for each participant for the trial.*

	Laboratory Number						
	1	3	4	5	6	7	8
(a) Laboratory Bias							
Calcium carbonate (C) (‰)	−1.01	1.16	−0.37	0.06	NR*	0.20	NR*
Caffeine (C) (‰)	−0.07	1.20	−0.61	−0.55	0.53	−0.41	0.09
Sucrose (C) (‰)	0.13	1.82	0.14	−0.23	0.02	−0.10	−0.41
Flour (C) (‰)	0.10	0.81	−0.16	−0.07	0.05	−0.25	0.03
Ammonium nitrate (N) (‰)	0.82	−0.04	−0.06	−0.17	0.24	−1.28	0.10
Potassium nitrate (N) (‰)	−0.22	0.02	−0.02	0.07	0.36	−0.27	0.02
Ammonium thiocyanate (N) (‰)	0.44	−0.28	−0.20	−0.11	0.12	−0.10	0.26
Flour (N) (‰)	0.60	NR*	0.97	−1.17	NR*	−0.40	0.00
(b) z-score							
Calcium carbonate (C) (‰)	−1.26	1.44	−0.46	0.07	NR*	0.25	NR*
Caffeine (C) (‰)	−0.10	1.76	−0.89	−0.81	0.77	−0.60	0.13
Sucrose (C) (‰)	0.54	7.73	0.60	−0.97	0.09	−0.44	−1.74
Flour (C) (‰)	0.62	4.90	−0.99	−0.40	0.29	−1.52	0.16
Ammonium nitrate (N) (‰)	3.43	−0.15	−0.25	−0.69	0.99	−5.36	0.41
Potassium nitrate (N) (‰)	−1.96	0.14	−0.16	0.66	3.29	−2.50	0.15
Ammonium thiocyanate (N) (‰)	1.72	−1.07	−0.78	−0.42	0.46	−0.38	1.01
Flour (N) (‰)	0.63	NR*	1.01	−1.23	NR*	−0.42	0.00

*Indicates that no result/data was reported.

TABLE 9—*Summary of assigned value (X) and standard uncertainty of assigned value (μ_X) for each of the chemicals distributed during the trial.*

	Assigned Value, X (‰)	Standard Uncertainty of Assigned Value, μ_X (‰)
Calcium carbonate (C)	−42.5	0.5
Caffeine (C)	−40.3	0.3
Sucrose (C)	−11.4	0.1
Plain flour (C)	−22.5	0.1
Ammonium nitrate (N)	−0.1	0.1
Potassium nitrate (N)	−0.4	0.1
Ammonium thiocyanate (N)	3.03	0.1
Plain flour (N)	4.8	0.5

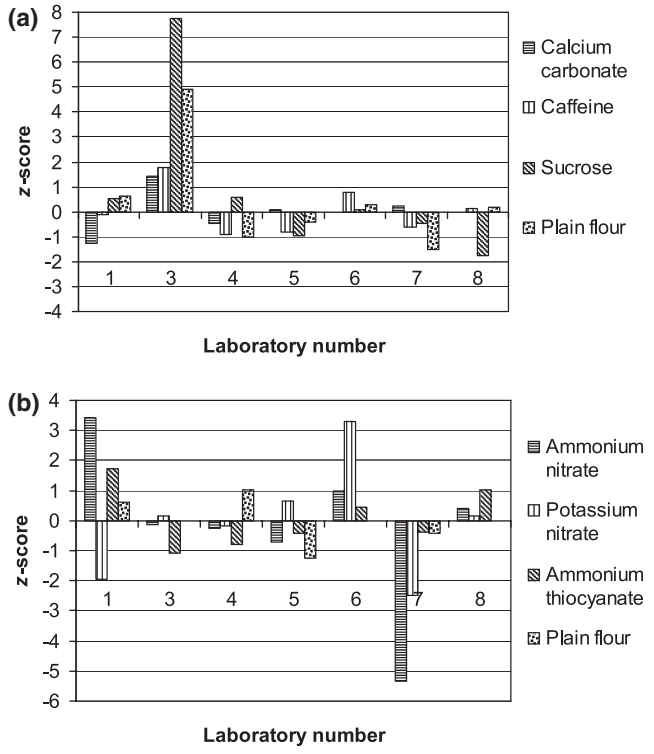

FIG. 1*(a)—Bar chart of* z*-scores for samples measured for carbon isotope. (b) Bar chart of* z*-scores for samples measured for nitrogen isotope.*

A second trial will evaluate whether any trends exist across the laboratories and whether implemented corrective actions were effective. The trial will also evaluate other chemical compounds of interest to forensic laboratories, e.g., potential laboratory standards for different target materials and chemicals displaying a wider isotopic range. Samples will be distributed to each of the participants as bulk samples requiring weighing and measuring (as conducted in this trial) but also as a tray of pre-weighed samples and standards (in the same range). The participants will be requested to measure the samples utilizing their standard procedures. This will assist in highlighting potential variation introduced through different sample preparation techniques. Consideration will be given to requesting participants to also measure the samples utilizing a specific method to assist in determining the true value of the samples and observing variations that may arise as a result of different instrument methods. Raw data will also be obtained from each participant and subsequently processed utilizing a common algorithm to highlight potential bias introduced through software calculation algorithms. Despite the need for future trials, the outcomes from this initial trial are extremely valuable as it is the first of its kind in the region and is essential for the implementation of IRMS by forensic science laboratories.

References

1. Benson S, Lennard C, Maynard P, Roux C. Forensic applications of isotope ratio mass spectrometry—a review. Forensic Sci Int 2006;157(1):1–22.
2. Kipphardt H. Traceability in isotopic measurements. In: de Groot PA, editor. Handbook of stable isotope analytical techniques. Vol. 1. Amsterdam: Elsevier B.V, 2004;928–43.
3. http://www.forensic-isotopes.rdg.ac.uk/newslett/newsletters/issue5.pdf.
4. http://www.forensic-isotopes.rdg.ac.uk/newslett/newsletters/issue7.pdf.
5. http://www.forensic-isotopes.rdg.ac.uk/newslett/newsletters/issue8.pdf.
6. International Standard ISO 13528 Statistical methods for use in proficiency testing by interlaboratory comparisons. ISO, First Edition 2005-09-01. Reference number: ISO 13528:2005(E), http://www.iso.org/iso/home.htm
7. Coplen T, Brand W, Gehre M, Gröning M, Meijer H, Toman B, et al. New guidelines for $\delta^{13}C$ measurements. Anal Chem 2006;78:2439–41.

Additional information and reprint requests:
Sarah Jane Benson, Ph.D.
Australian Federal Police
GPO Box 401
Canberra 2601 ACT
Australia
E-mail: sarah.benson@afp.gov.au

Appendix—Algorithm Utilized for Determining the Assigned Value, Standard Deviation for Proficiency Assessment, Standard Uncertainty of the Assigned Value and Other Performance Statistics

The assigned value and standard deviation for proficiency assessment were calculated using the following algorithm as detailed in ISO 13528:

1. List each participant's reported average (x_1, x_2, x_3 …. x_i)
2. Calculate the initial robust average (x^*) by calculating the median of the reported averages
3. Calculate the initial robust standard deviation (s^*) using the following formula:

$$s^* = 1.483 \text{ median of } |x_1 - x^*|$$

In order to calculate this:

a. calculate $|x_1 - x^*|$ for each participant
b. calculate the median of the values calculated in step a
c. multiply the median by 1.483

4. Calculate:

$$\delta = 1.5s^*$$

5. Calculate x_i^* for each of the participant's reported averages:

$$x_i^* = \begin{cases} x^* - \delta, & \text{if } x_i < x^* - \delta \\ x^* + \delta, & \text{if } x_i > x^* + \delta \\ x_i, \text{otherwise} & \end{cases}$$

Use δ as calculated in step 4; x_i is the average value initially reported by each participant and x^* is the initial robust average calculated in step 2

6. Calculate the final robust average using the following formula:

$$x^* = \Sigma x_i^* / p$$

x_i^* as calculated for each participant in step 5; p is the total number of participants

7. Calculate the final robust standard deviation using the following formula:

$$s^* = 1.134\sqrt{\Sigma(x_i^* - x^*)^2/(p-1)}$$

x_i* as calculated in step 5 for each participant
x* as calculated in step 6
p = number of participants

8. Use the following formula to determine the standard uncertainty of the assigned value:

$$\mu_x = 1.25 \times s^*/\sqrt{p}$$

s* is the robust standard deviation (for proficiency assessment) calculated in step 7

9. Laboratory bias (D) was estimated using:

$$D = x - X$$

x = average result reported from each participant
X = assigned value

10. z-scores were determined using:

$$z = (x - X)/\sigma$$

x = average result reported from each participant
X = assigned value
σ (or s*) = standard deviation for proficiency assessment as calculated in step 7