ARTICLE

JAAS
www.rsc.org/jaas

Improved calibration strategy for measurement of trace elements in biological and clinical whole blood reference materials *via* collision-cell inductively coupled plasma mass spectrometry

Steven J. Christopher,*[a] Russell D. Day,[a] Colleen E. Bryan[a] and Gregory C. Turk[b]

[a] *National Institute of Standards and Technology, Chemical Science and Technology Laboratory, Analytical Chemistry Division, 331 Fort Johnson Road, Charleston, SC 29412, USA. E-mail: steven.christopher@nist.gov; Fax: (843) 762-8742; Tel: (843) 762-8856*

[b] *National Institute of Standards and Technology, Chemical Science and Technology Laboratory, Analytical Chemistry Division, 100 Bureau Drive, Gaithersburg, MD 20899, USA*

Received 22nd March 2005, Accepted 1st August 2005
First published as an Advance Article on the web 18th August 2005

A multi-element quantification strategy based on the method of standard additions incorporating internal standards and collision cell inductively coupled plasma mass spectrometry (ICP-MS) is presented. Approaches to experimental design are discussed in the context of streamlining analytical measurement protocols employing ratio-based standard additions quantification schemes for the certification of multiple elements in Certified Reference Materials, including reduction of the number of required analytical samples and measurement of analytes and internal standards at alternate quadrupole mass resolution settings. This strategy was implemented for the measurement of As, Se, Fe, Mn, Rb, Cu, and Zn levels in a candidate fish tissue NIST Standard Reference Material and measurement of Cd and Pb in two clinical, whole blood Certified Reference Materials. A simple approach to calculating analytical uncertainties for concentration data, as determined using standard additions calibrations, is presented which utilizes regression and prediction uncertainties and quotient propagation of error formulae.

Introduction

Measurements performed by National Metrology Institutes such as the National Institute of Standards and Technology (NIST) to assign the certified values to reference materials are a unique and important aspect of analytical chemistry. For natural-matrix Certified Reference Materials (CRMs) the process typically begins with methods designed to test the homogeneity of a candidate material, followed by high-accuracy value assignment measurements for the analytes of interest. Often, one definitive analytical technique (*e.g.* isotope dilution mass spectrometry) is used for value assignment measurements. Alternatively, two or more independent methods of analysis are used to assess and account for analytical method-associated systematic biases. At NIST, this data is generated both internally and externally through the vehicle of round-robin exercises and coordination of or participation in other quality assurance activities. Data are compiled from various sources, statistically evaluated, and combined to arrive at a mean and expanded uncertainty about the mean. This data becomes the end product published on an official Certificate of Analysis. One of the CRM materials examined in this study, candidate Standard Reference Material (SRM) 1947 Lake Michigan Fish Tissue, serves as a useful example to describe the processes outlined above. SRM 1947 Lake Michigan Fish Tissue was collected with the assistance of the Michigan Department of Natural Resources and is a fresh-frozen material prepared from fillets (79 kg) of adult lake trout (*Salvelinus namaycush*). The material was cryo-pulverized and homogenized at NIST and subsequently bottled for characterization and value assignment of various proximates and organic and trace element constituents. The discussion forthwith focuses solely on the trace element component of the value assignment process for SRM 1947 Lake Michigan Fish Tissue. First the material was characterized for homogeneity at NIST, and subsequently disseminated to approximately 20 laboratories in 2000, through a round-robin exercise coordinated by the Food Processors Association and to approximately 30 laboratories in 2001, through the NIST-coordinated Interlaboratory Comparison Exercise for Trace Elements in Marine Mammals, sponsored by the National Marine Fisheries Service. This data, along with NIST data, including that derived from the methods of analysis presented in this study, comprises the large body of work that will eventually be used to assign trace element concentration values to SRM 1947.

Here we describe a value assignment measurement procedure recently developed at NIST for the determination of trace elements in NIST Standard Reference Materials (SRMs) using inductively coupled plasma mass spectrometry (ICP-MS). We apply this procedure to two example experiments that are representative of the research we routinely perform in our laboratory. The first example is a certification experiment where trace element concentrations are to be assigned to a candidate SRM, in this case SRM 1947 Lake Michigan Fish Tissue, using the method of standard additions. The second example is a clinical experiment where we are validating the accuracy of a new standard additions method for the determination of trace elements in whole blood samples, using whole blood certified reference materials, before application to a high sample load project requiring implementation of a high-accuracy, matrix-matched calibration scheme. Accuracy is the paramount concern for such measurements, and for this reason isotope dilution procedures are often used at NIST whenever possible and practical.[1] However, for projects where we need to simultaneously measure both multiisotopic and monoisotopic trace element analytes accurately, complementary standard addition methods are often developed and applied. The method described here is a high-accuracy, multi-element procedure

DOI: 10.1039/b504139b

complementary to isotope dilution that can also be applied to the quantification of both multiisotopic and monoisotopic elements. The method is based on the use of standard additions combined with the use of internal standards and gravimetric solution handling. Though originally developed for trace element certification metrology (typically a low sample throughput endeavor), we have found that this procedure is simple enough to be routinely applied to high sample load measurement problems that require a close matrix match between samples and calibration standards.

In most laboratories, the method of standard additions is implemented using volumetric sample handling for solution-based methods such as ICP-MS. Samples are mixed with varying amounts of the standard addition spike in volumetric flasks, and all solutions are brought to a constant volume. This gives the matrix-matched conditions relied upon for the elimination of multiplicative matrix effects and compensates for dilution that occurs when a spike solution is added to a sample solution. We prefer to use gravimetric rather than volumetric solution handling, for improved accuracy and reliability. However, bringing sets of solutions to a constant mixed mass is tedious to the point of being impractical. This necessity can be avoided by using an internal standard. Here, the dilution that occurs when a standard addition spike is added affects the internal standard exactly as it affects the analyte, and the ratio is not affected. Spiked and unspiked solutions can be diluted to approximately the same mass, and the internal standard will compensate for slight variations.

The primary use of internal standardization in ICP-MS is to correct for multiplicative matrix effects from sample to sample where signal suppression or enhancement occurs.[2,3] We typically use internal standards even when quantifying by the method of standard additions, but it is not a prerequisite for obtaining accurate data,[4,5] as when employing external calibration approaches. For the last several years we have used a "split and spike" standard additions procedure. In this procedure a single sample is dissolved (if a solid), spiked with an internal standard, and gravimetrically split into 2 portions: one portion is then spiked with the standard addition while the other remains unspiked. This procedure assures a matched internal standard concentration in the split portions, but requires extensive quantitative post-digestion sample handling, especially considering that multiple elements of varied concentrations are being measured and multiple SRM units are being analyzed. Standard addition calculations are then based on ratios of analyte to internal standard, which tend to improve precision and partially correct for instrumental drift.[6] Recently we have modified this approach for certification experiments where multiple units of a NIST SRM need to be analyzed. The new standard additions/internal standard procedure described here takes advantage of the fact that we are analyzing multiple samples of a homogeneous CRM. In the new procedure, multiple (typically 6) samples of the candidate CRM are accurately weighed into digestion vessels and gravimetrically spiked with an internal standard solution. Varying amounts of a multi-element standard additions spike solution are then gravimetrically added to each vessel, ranging from no spike to a high spike, ideally resulting in a three- to four-fold analyte increase. The samples are then digested. From this point on, all measurements will be based on analyte/internal standard ratios, and all sample handling, including removal of the solutions from the digestion vessels and any subsequent dilutions, does not need to be done quantitatively. This is one of the main benefits of this approach. The other benefit is that for homogeneous samples we do not have to perform sample splits or multiple standard additions on each sample (unit) to be measured. Again, in terms of calibration this latter benefit is realized because we are dealing with well-characterized, homogeneous samples of identical matrix. However, the benefit propagates to unspiked unknown and unspiked control materials run against the calibration curve so long as the matrices remain similar to the calibrants, but not necessarily identical. In this study for example, we sucessfully ran complementary fish tissue and blood control samples against calibration curves generated using similar tissues. Values and uncertainties are calculated by linear regression of the measured analyte/internal standard ratios *versus* the amount of standard added, with appropriate compensation for variability between sample masses and internal standard spikes.

The two-fold purpose of the work presented here was to develop a more streamlined approach to multi-element standard additions for ICP-MS as applied to the certification of NIST Standard Reference Materials, and also to develop methods to analyze critical marine biological and clinical samples with high sample load demands while preserving matrix-matched calibration schemes. The present study focuses on a limited suite of elements and a single mode of collision cell operation at two different quadrupole mass resolution settings: however, the quantification strategy presented could be further streamlined by taking advantage of the available technology that collision cell ICP-MS instrumentation provides—near simultaneous, multimode analysis capabilities, including normal mode and collision cell modes of operation using different collision gases, at variable instrument and mass resolution settings. The software adjustable mass resolution feature built into the quadrupole ICP-MS instrument used in this study provides element/isotope specific, real-time, switchable instrument mass resolution settings,[7] where unit and high resolution (0.2 u) measurement capabilities assist with the determination, respectively, of low and high concentration analytes of interest in the sample. This greatly aids in simplifying the experimental handling of solutions destined for multi-element analysis because it reduces the number of separate portions of sample that need to be prepared, appropriately diluted, and analyzed.

Experimental methods

Disclaimer

Certain commercial products are identified in this paper to adequately specify the procedures outlined. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology.

Instrumentation

A ThermoElemental X-7 inductively coupled plasma mass spectrometer (ThermoElemental, Franklin, MA) was used for all of the experiments. The ICP-MS instrument has the capability to perform acquisitions at unit or high mass resolution in both normal and collision cell modes. Here the term high resolution is defined by the instrument manufacturer relative to the unit resolution of a typical quadrupole instrument operating in the first stability region and is not in any way comparable to the high resolution term or capabilities ascribed to mass spectrometers that employ sector field technology. High resolution is effectively used in this quadrupole unit to attenuate the ion beam (not eliminate isobaric interferences) for measurements involving elements at high (single-mg kg^{-1}) concentrations.[7] Introduction of a collision gas into a hexapole or quadrupole ion guide prior to mass filtering with a standard quadrupole unit allows for the destruction of meddlesome, plasma-based isobaric interferences.[8] For the certification experiment example presented, the high resolution, collision cell mode was used for Fe and Rb in the samples, while the remaining elements (As, Cu, Mn, Se, and Zn) were measured in collision cell mode at unit mass resolution. Most of these elements can potentially suffer from isobaric interferences induced by either the ICP working gas (Ar) alone or in combination with elements emanating from the sample and

nitric acid matrix. For the clinical experiment example presented, low and high resolution collision cell modes, respectively, were used for measurement of Cd and Pb. The typical ICP-MS instrument settings used for both experiments are tabulated in Table 1. Most elements were measured using a 60 s acquisition with peak jumping and dwell times ranging from 20 ms to 50 ms, depending on isotopic abundance and element concentration. Potentials for the ion optics and quadrupole were read directly from software. The high purity collision gas mixture employed was 8% H_2 in He operating at either a flow rate of approximately 7 mL min^{-1} (certification experiment) or 8.5 mL min^{-1} (clinical sample experiment). Gas flows for the collision cell were metered using instrument software in combination with a system mass flow controller calibrated for pure He gas. There is no direct way of measuring the collision cell gas pressures in the hexapole cell, but the following pressures are given to inform the reader of the typical conditions encountered in the study: base pressure of instrument $\sim 8 \times 10^{-8}$ mBar, analyzer pressure under ICP gas load without collision cell $\sim 4 \times 10^{-7}$ mbar, analyzer pressure under ICP gas load with 8.5 mL min^{-1} of collision gas flow $\sim 2 \times 10^{-6}$ mbar. A tuning solution containing 10 ng g^{-1} of Co, 10 ng g^{-1} of In and 20 ng g^{-1} of Pb in 2% HCl (mass fraction) was used to tune the system. This consisted of optimizing signal intensity and stability at m/z 59 (Co), 115 (In) and 208 (Pb), while carefully monitoring and minimizing (<200 counts s^{-1}) the

Table 1 Instrument parameters for certification and clinical experiments

Isotopes monitored		Quadrupole ICP-MS settings	
Measured elements	Internal standard reference	Dwell times (ms) isotope(s)/internal standard	Resolution
Certification experiment			
^{75}As	^{99}Ru	30/20	Unit
$^{63,65}Cu$	^{99}Ru	20,20/20	Unit
	^{103}Rh,	50,25/25	
	^{101}Ru	/20	
$^{54,56}Fe$	^{102}Ru	/20	0.2 u
	^{115}In	/20	
	^{93}Nb	/20	
^{55}Mn	^{99}Ru	20/20	Unit
^{85}Rb	^{103}Rh	25/25	0.2 u
^{80}Se	^{99}Ru	30/20	Unit
$^{64,66,68}Zn$	^{99}Ru	40,25,25/20	Unit
Ion optics (V)			
Extraction	Lens 1	Lens 2	Lens 3
−481	−0.20	−50.3	−96.4
Deflector 1	Focus	Deflector 2	Differential aperture
−52.6	+9.2	−175	−148
Quadrupole bias	Hexapole bias		
−3.0	+0.20		
Torch settings			
Power (W)	Cool gas (L min^{-1})	Auxiliary gas (L min^{-1})	Nebulizer gas (L min^{-1})
1200	13.0	0.96	0.90
Collision gas settings			
H_2/He volume fraction (%)	Gas flow (mL min^{-1})		
8.0% H_2	7.0		
92% He			

Isotopes monitored		Quadrupole ICP-MS settings	
Measured elements	Internal standard reference	Dwell times (ms) isotope(s)/internal standard	Resolution
Clinical experiment			
^{111}Cd	^{89}Y	20/20	Unit
$^{206,207,208}Pb$		20,20,20	0.2 u only
	^{103}Rh	/25	0.2 u
	^{115}In	/20	0.2 u
	^{197}Au	/20	Unit
	^{141}Pr	/20	Unit
Ion optics (V)			
Extraction	Lens 1	Lens 2	Lens 3
−381	+4.8	−50.3	−96.4
Deflector 1	Focus	Deflector 2	Differential aperture
−52.6	+11.2	−175	−148
Quadrupole bias	Hexapole bias		
−3.0	+0.20		
Torch settings			
Power (W)	Cool gas (L min^{-1})	Auxiliary gas (L min^{-1})	Nebulizer gas (L min^{-1})
1200	13.0	0.96	0.94
Collision gas settings			
H_2/He volume fraction (%)	Gas flow (mL min^{-1})		
8.0% H_2	8.5		
92% He			

background contributions at m/z 56 ($^{40}Ar^{16}O^+$), m/z 75 ($^{40}Ar^{35}Cl^+$), and m/z 80 ($^{40}Ar_2^+$). Thereafter, a solution containing the elements of interest was introduced into the system and short term stability measurements were performed over 10 min (ten 60 s acquisitions) to ensure that the system was operating satisfactorily in terms of element sensitivity and stability (<1% RSD over 10 min).

Sample preparation

Safety warning. Certain aspects of the sample preparation scheme require working with strong oxidizing acids under conditions of elevated temperature and pressure, ammonium hydroxide, 1-butanol, and human and animal blood, all of which require the use of chemical fume hoods, specialized personal protective equipment and (for blood) biohazard handling procedures.

SRM certification experiment. To test the new multi-element standard additions certification method, six units of candidate Standard Reference Material SRM 1947 Lake Michigan Fish Tissue were selected from cases from two large production batches of SRM 1947 using a random stratified sampling scheme. SRM 1946 Lake Superior Fish Tissue served as a control in the exercise, possessing similar matrix properties (both samples are comprised of edible fish tissue from lake trout) and similar element concentrations. The sample dissolution procedures used closed vessel microwave digestion. For a twelve-sample digestion batch, approximately 1 g sample aliquots from each unit of SRM 1947 ($n = 6$) were digested along with two 1 g control sample aliquots (SRM 1946) and four procedural blanks. Sample, internal standard, and analytical spike aliquots delivered to the digestion vessels were weighed by difference on a three-decimal place balance with 0.001 g readability. All samples were weighed directly into tared microwave decomposition vessels and spiked with a 1 g aliquot of a multi-element internal standard solution, and additionally the SRM 1947 samples were spiked with 0–1 g aliquots of a multi-element spike solution containing a mass fraction of 50 : 50 high purity water and concentrated nitric acid as the diluent. Proportions of elements in the multi-element spike solution were derived from concentration values obtained from a NIST/NOAA Interlaboratory Comparison Exercise in 2001, where SRM 1947 was issued as an unknown to 30 laboratories.[9] Various aliquots of make-up nitric acid were added to each sample, including the SRM 1946 control samples and blanks, to account for the different acid composition generated from the standard additions spiking procedure for SRM 1947, in order to matrix-match the samples. Samples were digested in a medium of approximately 5.5 g of high-purity nitric acid in a PerkinElmer Multiwave® (PerkinElmer Corporation, Norwalk, CT, USA) microwave oven operating at the highest possible temperatures (up to 300 °C) and pressures (up to 8000 kPa). The resulting digests were vented and diluted with high purity water to a total volume of approximately 50 mL and non-quantitatively transferred into precleaned (dilute nitric acid soak) 50 mL low-density polyethylene centrifuge tubes.

Clinical sample experiment. A standard additions calibration scheme similar to the one described in the certification experiment was employed, but without the assistance of microwave sample decomposition. All clinical samples were prepared in a diluent consisting of several high-purity chemicals: 1% 1-butanol + 1% ammonium hydroxide + 0.01% EDTA + 1% Triton X-100 + 1% nitric acid in balance high-purity water (mass fraction). The combined effects of this chemical suite serve to reduce the viscosity of the nebulized samples and ameliorate the differential ionization effects induced by carbon in individual analytical samples. 1-Butanol acts to reduce matrix effects by elevating the carbon in all matrices so that variable, sample-to-sample ionization effects are minimized, especially for As and Se.[10,11] Sieniawska *et al.* in ref. 12 have discussed how the suite of chemicals used here helps to prepare the whole blood samples for ICP-MS determination: ammonium hydroxide lyses red blood cells, EDTA prevents the loss of metals by precipitation or absorption, and Triton X-100 prevents blockage of the ICP-MS nebulizer and torch injector tube. After sample solubilization, multi-element internal standards and spikes were also added to each sample, where appropriate. Finally, 1% high-purity nitric acid was added to stabilize the trace elements in each sample. The end solution resulted in a 1 : 30 dilution of whole blood carried out in precleaned screw cap 68 mL polypropylene (autosampler) tubes. The order in which the blood sample is mixed is crucial. One must avoid adding acid directly to the blood (*e.g.*, from the internal standard solution, standard addition spike solution, *etc.*) before the diluent, as the blood will coagulate, resulting in a particle-laden, inhomogeneous sample. The mixing scheme allowed samples to be run in a "dilute and shoot" manner (without a high temperature and pressure sample decomposition).

Calibration methods and sample measurements

Prior to digestion, the SRM 1947 samples were spiked with a 1 g aliquot of a multi-element internal standard solution containing Nb, Y, Ru, Rh, In, Pr, and Au, and each analytical sample was spiked with increasing aliquots of a multi-element spike solution containing As, Cu, Fe, Mn, Rb, Se, and Zn to create a set of standard additions calibration curves. SRM 1946 control samples and the procedural blanks were spiked with internal standard elements similar to SRM 1947. Samples were run in a random order in a block of 5 repeat sequences to capture the effect of instrumental drift. The measured analyte to internal standard ratios (R) for most elements were corrected for instrumental drift using a third order polynomial fit (As, Cu, Fe, Mn, Rh, and Se), except for Zn, which exhibited first order drift. The slope and intercept of the calibration curves (plot of $Y = R \times$ gram of internal standard per gram of sample as a function of $X =$ micrograms of analyte spike added per gram of sample, with the response on the Y axis blank corrected prior to plotting) were used to calculate the concentration of each element in SRM 1947. The slopes resulting from the SRM 1947 calibration curves and the measured analyte responses from the SRM 1946 and method blank samples were used to determine element concentrations in the SRM 1946 control samples. A similar calibration approach was utilized when studying the clinical samples. Uncertainty budget calculations for the experiments in this study consider only the uncertainty due to regression and prediction, where appropriate, as these factors far outweigh other sources of uncertainty due to drift, spike calibrants, blanks and internal precision measurements.

Expanded uncertainties for the trace element concentration data presented in the study are derived using quotient propagation of error formulae as presented in the next section.

Results and discussion

SRM certification experiment

The goal of an analyte value assignment measurement for CRM certification is to accurately assign the certified value to a homogeneous lot of candidate CRM that may include several hundred CRM units, and to accurately assess the uncertainty of the measurement. The procedures described here help streamline this task. Multiple units of SRM 1947 were statistically sampled from the entire lot and a standard additions experiment was designed with *a priori* knowledge of the rough analyte concentrations prior to sample analysis. Pre-existing knowledge of the concentration is typical for any

CRM that has been disseminated through some type of round-robin or proficiency testing effort. Spiking each sample with internal standard and an analytical spike aliquot prior to digestion or equilibration eliminates the need for quantitative handling of the digests as long as the analyte/internal standard ratio is taken as the response for the standard additions curve. Although the analyte/internal standard ratio is essentially fixed after sample decomposition, carefree dilution of the analytical samples should be avoided, in order to maintain a good matrix match between analytical standards, controls, and unknown samples. Spiking the SRM 1947 samples with varying levels of standard across the representative batch of SRM units eliminates the need to perform sample splits, which reduces the number of samples significantly. Utilizing both the unit and high resolution modes of the ICP-MS allows for multi-element analysis of a single set of analytical samples and standards. Finally, incorporating internal standards and correcting for multiplicative drift using correction procedures developed at NIST[13] can improve the precision and accuracy dramatically.

The calibration curves generated for the certification experiment should back cast to the mean concentration at the *X*-intercept for the representative batch of SRM 1947 samples and provide good estimates of the true concentration values of elements present in the material. The slope of such a plot (and the response of the control sample) can then be used to verify accuracy by predicting (determining) the concentration in a similar-matrix control material (SRM 1946 Lake Superior Fish Tissue in this example). Analytical uncertainty due to regression for the target material and analytical uncertainty due to prediction for the control material can readily be obtained. These errors cast a window of uncertainty about the point of extrapolation (estimated concentration) at the *X*-intercept. Fig. 1 shows an example calibration curve from the certification experiment for Cu in SRM 1947. The large gap in the calibration curve was purposely designed into the experiment and is reflective of how we historically performed our single point "split and spike" standard additions experiments (approximately 3× increase in signal between the zero addition and spiked sample) to minimize *X*-intercept extrapolation uncertainty induced by the calculated slope; if the samples are too underspiked in a single point standard addition, this will serve to reduce the slope and the uncertainty about each calibration point will cast a greater range of possible values about the *X*-intercept. Our multi-element results (see Tables 3 and 4) show that the gap in the calibration curve does not significantly impact the accuracy of the experiment, but in hindsight we would prefer to have a more evenly distributed multi-point standard additions calibration curve to avoid leverage effects that may potentially contribute to extrapolation errors. This situation was remedied when performing calibrations for the clinical sample experiment. The measured concentration of Cu in SRM 1947 is simply calculated as the quotient of the regression parameters of intercept and slope, and the measured concentration in the control material is calculated by dividing the response of the SRM 1946 sample by the slope calculated from the SRM 1947 calibration curve. Table 2 lists the corresponding regression output statistics for the ^{63}Cu calibration curve, calculated from the LINEST function in Microsoft Excel (Microsoft, Seattle, WA, USA). This table is put forward to serve as an example of how the variance parameters are obtained and used to calculate the uncertainty about the *X*-intercept (uncertainty for the estimated concentration) for a standard additions calibration curve. Regression uncertainties for various trace elements in SRM 1947 were determined using the quotient propagation of error formula to estimate a variance at the *X*-intercept ($y = 0$) for a linear calibration fit. The formula for regression uncertainty for SRM 1947 is:

$$\mathrm{Var}\left[\frac{-b}{m}\right] = \left[\frac{b^2}{m^2}\right] \times \left[\frac{V_b}{b^2} + \frac{V_m}{m^2}\right] \qquad (1)$$

where $b = Y$-intercept, $m =$ slope of calibration curve, and the V_b and V_m terms refer to the squared uncertainty estimates for the intercept and slope, respectively. V_b and V_m can readily be obtained using the uncertainty estimates calculated from the LINEST function (see example for ^{63}Cu in Table 2). The control sample (SRM 1946) uncertainty incorporates a prediction term along with the slope and intercept terms outlined in eqn. (1):

$$\mathrm{Var}\left[\frac{y_0 - b}{m}\right] = \left[\frac{(y_0 - b)^2}{m^2}\right] \times \left[\frac{V_b + V_y}{(y_0 - b)^2} + \frac{V_m}{m^2}\right] \qquad (2)$$

where $b = Y$-intercept, $m =$ slope of calibration curve, and as above, the V_b and V_m terms refer to the squared uncertainty estimates for the intercept and slope, respectively. The y_0 and V_y terms refer to the measured response for the control sample (SRM 1946) and its associated prediction uncertainty, respectively. V_y (see Table 2) can also be obtained from the LINEST error estimates and is defined as the square of the standard uncertainty of the predicted *y*-value for each *x* in the regression. For a standard additions experiment, only the slope of the calibration curve is used to estimate the element concentration in the control sample or an unknown sample that is run against the calibration curve, so eqn. (2) is evaluated with the b and V_b terms set equal to zero to result in the prediction uncertainty of the control (or unknown) sample (eqn. (3)):

$$\mathrm{Var}\left[\frac{y_0}{m}\right] = \left[\frac{y_0^2}{m^2}\right] \times \left[\frac{V_y}{y_0^2} + \frac{V_m}{m^2}\right] \qquad (3)$$

Table 3 lists the measured results and expanded uncertainties for various elements in SRM 1947 along with the consensus mean data and expanded uncertainties obtained from the 2001 Interlaboratory Comparison Exercise for Trace Elements in Marine Mammals. The interlaboratory comparison data is

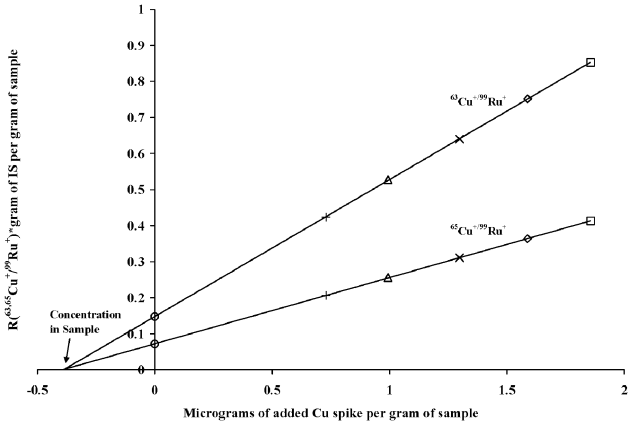

Fig. 1 Example standard additions calibration curves for $^{63,65}Cu^+$ isotopes in SRM 1947 Lake Michigan Fish Tissue referenced to $^{99}Ru^+$ internal standard isotope. Symbols represent sample aliquots from six distinct units of material spiked at different concentration levels. Calibration equations: ^{63}Cu, $y = 0.000\,38x + 0.147\,81$, $R^2 = 0.999\,98$; ^{65}Cu, $y = 0.000\,18x + 0.072\,43$, $R^2 = 0.999\,97$.

Table 2 Example linear regression output for ^{63}Cu in candidate SRM 1947 Lake Michigan Fish tissue

Slope	Intercept
0.000 38	0.147 81
Standard uncertainty of the slope, $\sqrt{V_m}$	Standard uncertainty of the intercept, $\sqrt{V_b}$
8.96E-07	1.11E-03
Correlation coefficient	Standard uncertainty of prediction, $\sqrt{V_y}$
0.999 98	1.33E-03

Table 3 Summary of results for candidate SRM 1947 Lake Michigan fish tissue obtained using ratio-based standard additions calibration scheme and comparison to round-robin data from the 2001 NIST/NOAA interlaboratory comparison exercise for trace elements in marine mammals

Results using ratio-based standard additions calibration scheme			Consensus results from the 2001 NIST/NOAA interlaboratory comparison exercise		
Element	Measured value (mg kg^{-1})	Expanded uncertainty (mg kg^{-1})	Number of laboratories (SRM units)	Consensus mean (mg kg^{-1})	Expanded uncertainty (mg kg^{-1})
As	0.752	0.062	16	0.717	0.042
Cu	0.389	0.008	19	0.421	0.029
Fe	3.48	0.38	21	3.72	0.36
Mn	0.073	0.004	16	0.092	0.014
Rb	4.49	0.19	14	4.54	0.10
Se	0.491	0.055	18	0.441	0.045
Zn	2.67	0.25	22	2.63	0.15

derived from application of a maximum likelihood consensus mean estimator algorithm developed at NIST by Rukhin and Vangel.[14] There is good agreement between the datasets. The measured values from SRM 1946 predicted by the SRM 1947 standard additions calibration curves and expanded uncertainties are compared with certified values for SRM 1946 in Table 4. The consensus mean data and the SRM 1946 control data presented in Tables 3 and 4, respectively, validate the accuracy of the calibration scheme employed here.

The quality (in terms of correlation coefficient values and *X*-intercept extrapolation precision) of the standard additions calibration curves was assessed for several elements using different internal standards as points of reference. In general, most internal standards yielded good linear calibration curves. Example calibration curves for As and Fe, measured in unit and high resolution modes, respectively, are presented in Figs. 2 and 3. Calibration curves were generated for As using ^{99}Ru and ^{89}Y as internal standards and Fe calibration curves were generated using ^{93}Nb, ^{101,102}Ru, ^{103}Rh, and ^{115}In as internal standards. Calibration curves for As converge within a $\pm$0.3% window at the *X*-intercept and demonstrate reasonable extrapolation precision, regardless of the internal standard point of reference, with an average correlation coefficient of 0.9996 $\pm$ 0.006% relative standard deviation. Residuals for both of the As calibration curves showed no trend as a function of concentration and ranged from $-$1.15% to +0.74% of the magnitudes of the response variables for As/Ru and from $-$1.28% to +0.74% of the magnitudes of the response variables for the As/Y calibration curves. Similar results were observed for the Fe calibration curves and residuals (average correlation coefficient of 0.9996 $\pm$ 0.012% relative standard deviation). A closer look at the *X*-intercept region for all of the Fe calibration curves (Fig. 4) shows convergence of the calibration curves over roughly a $\pm$3% window, with the subset of data for the Rh and Ru internal standards forming a tighter point of convergence ($<\pm$0.5% window). A comparison of first ionization potentials (IPs) for the Fe analyte (7.87 eV) with IPs of internal standards for In (5.79 eV), Nb (6.88 eV), Ru (7.37), and Rh (7.46 eV) suggests that the constant matrix system can tolerate internal standards over a range of masses and ionization potentials. This is further exemplified by the calibration data for As presented in Fig. 2, where the large differences in IPs that exist between analyte (9.81 eV) and internal standards Y (6.38 eV) or Ru (7.37 eV) do not seem to adversely impact extrapolation precision or the correlation coefficients of the calibration curves. The examples presented above should be considered as representative of observations for most of the elements and internal standards studied. We did notice a possible weak trend in the intercept as a function of internal standard mass for the Fe data in Fig. 4, but it is unclear if this trend is real (perhaps an artifact of high resolution mode operation) when one considers the uncertainty about each of the intercepts. We also observed a slight curvature in the Se calibration curves when calculations were performed using ^{89}Y as the internal standard. The resulting calibration curves showed a general tendency to "roll over" at the highest spike levels, indicating a slight suppression in the ^{80}Se/^{89}Y ratio. This, of course, will impact on the slope (accuracy) of the calibration curve and the measurement of Se in any unknown or control samples run against the calibration curve. Referencing ^{99}Ru as the internal standard removes this effect, resulting in a better linear fit through the entire spike range. The subtle effects observed for Se are likely dependent on a number of factors, including the specific analyte/internal standard pair, the sample matrix, element homogeneity and the analytical instrument (collision gas effects, non-correlated drift, *etc.*). Further experiments would be necessary to isolate the responsible component but that is beyond the scope of this research. This exercise demonstrates that it is always prudent to comprehensively look at the analytical measurement system from the perspective of employment of different internal standards and the factors that can affect these points of reference.

Clinical sample experiment

Similar procedures were tested to see if matrix-matched calibration curves employing the method of standard additions could be used to predict concentrations in clinical (blood) samples. The matrix tolerance of the calibration scheme for clinical whole blood samples of different type (human and bovine whole blood) was tested by generating three standard additions calibration curves in succession for ^{111}Cd$^+$ (^{89}Y$^+$

Table 4 Summary of measured and certified values for the control sample, SRM 1946 Lake Superior fish tissue

Element	Measured value (mg kg^{-1})	Expanded uncertainty (mg kg^{-1})	Certified value or [Reference value] (mg kg^{-1})	Expanded uncertainty (mg kg^{-1})
As	0.309	0.070	0.277	0.010
Cu	0.420	0.010	[0.476]	0.060
Fe	4.03	0.46	4.00	0.32
Mn	0.073	0.005	—	—
Rb	8.04	0.25	—	—
Se	0.546	0.066	[0.491]	0.043
Zn	3.33	0.30	[3.10]	0.18

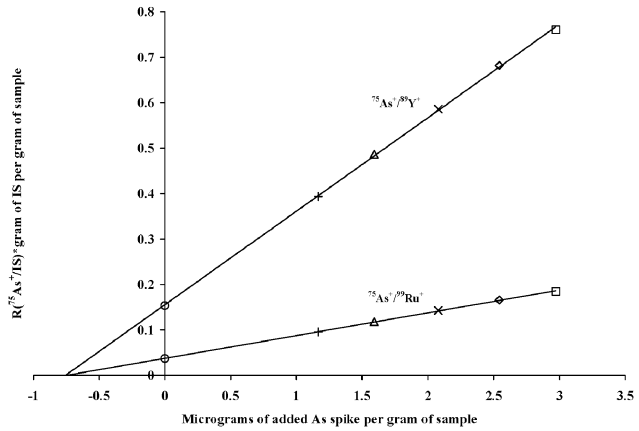

Fig. 2 Standard additions calibration curve for unit resolution monitoring of $^{75}As^{+}$ in SRM 1947 Lake Michigan Fish Tissue referenced to $^{89}Y^{+}$ and $^{99}Ru^{+}$ internal standard (IS) isotopes. Symbols represent sample aliquots from six distinct units of material spiked at different concentration levels.

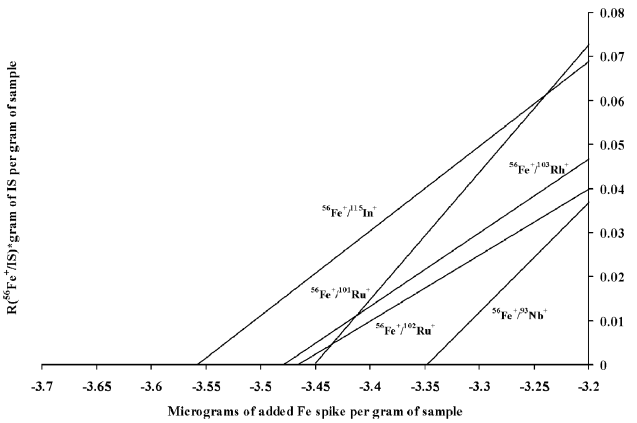

Fig. 4 Magnified *X*-intercept region for Fe calibration curves displayed in Fig. 3.

internal standard) and $^{206,207,208}Pb^{+}$ ($^{103}Rh^{+}$, $^{115}In^{+}$, $^{197}Au^{+}$, and $^{141}Pr^{+}$ internal standards) using several aliquots of a whole blood certified reference material, CRM Seronorm Trace Elements in Whole Blood Level III (batch No. 203056, Nycomed, Oslo, Norway). These data were not corrected for instrumental drift. This experiment differs from the certification example presented above. In this case, we are not using the calibration curves to back cast to the *X*-intercept to estimate a true value for a homogeneous batch of material (as in the certification example) but we are trying to establish the instrument response that will successfully predict element concentrations in subsequent samples that have similar matrix properties. One unique Seronorm sample and one unique SRM 966 sample were measured at the end of each calibration block. Thus, the slopes of the three calibration curves were used to predict the Cd concentration in three Seronorm samples and additionally in three samples of SRM 966, Toxic Elements in Bovine Blood, run as unknowns against the calibration curves. The average predicted concentrations and combined expanded uncertainties (simple 95% CL) for Cd in Seronorm (11.81 ng g^{-1} ± 1.88 ng g^{-1}) and SRM 966 (5.12 ng g^{-1} ± 0.84 ng g^{-1}) were in excellent agreement with density-corrected method reference values in Seronorm (sector field ICP-MS: 10.7 ng g^{-1} ± 0.6 ng g^{-1} and electrothermal atomic absorption spectrophotometry: 11.2 ng g^{-1} ± 0.8 ng g^{-1}) and density-corrected certified values in SRM 966 (4.95 ng g^{-1} ± 0.15 ng g^{-1}), respectively. Standard additions calibration curves using Seronorm were also generated for Pb, which is certified in Seronorm and SRM 966. The Pb results for Seronorm and SRM 966 samples calculated using the Seronorm regression data are presented in Fig. 5†. In the SRM certification example presented earlier, we used the same resolution mode for target analytes and internal standard points of reference. Here we test the robustness of ratio-based standard additions as a function of different internal standards run in both 0.2 u ($^{103}Rh^{+}$ and $^{115}In^{+}$) and 1.0 u ($^{197}Au^{+}$ and $^{141}Pr^{+}$) resolution modes, while measuring the target Pb isotopes only at a resolution of 0.2 u. It might be possible to reference analyte signal to internal standard signal regardless of the resolution setting used to measure either signal because the collision cell, ion optics and torch settings do not explicitly have to change when the mass resolution is altered. There are several instances where an analyst could find this feature useful, including, but not restricted to, reduction of the number of internal standards in the analytical system, for high resolution measurements that may suffer in terms of precision and in cases where different isotopes of an element are measured at different resolution due to gross isotopic abundance differences. However, a few caveats should be considered when the quadrupole resolution is different for each component of an analyte/internal standard pair. The peak shape of a quadrupole operating at unit resolution is not ideal for high precision measurements (*i.e.*, ratio measurements requiring less than ~0.1% relative standard deviation, for example). Under quadrupole operation at higher resolutions this problem would be further exacerbated. This scenario is borne out in Fig. 5—the Pb values obtained for Seronorm and SRM 966 using Seronorm standard additions calibration curves are in good agreement with the certified values, regardless of the resolution mode setting used to measure the internal standards. However, it is interesting to note that the Pb uncertainties† are slightly lower when internal standards are measured at unit resolution. It is hard to gauge from this limited dataset the utility of measuring internal standards at a mass resolution setting that

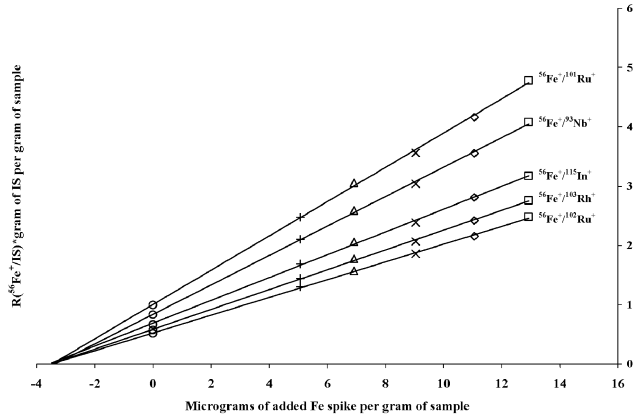

Fig. 3 Standard additions calibration curves for high resolution monitoring of $^{56}Fe^{+}$ in SRM 1947 Lake Michigan Fish Tissue referenced to $^{93}Nb^{+}$, $^{101,102}Ru^{+}$, $^{103}Rh^{+}$, and $^{115}In^{+}$ internal standard (IS) isotopes. Symbols represent sample aliquots from six distinct units of material spiked at different concentration levels.

† Note: Uncertainties for the Pb/internal standards mass fraction data are expressed at the 95% confidence level $\left(\frac{t \times s}{\sqrt{n}}\right)$, with $n = 3$ and df = 2. The Certificate of Analysis for Seronorm Trace Elements in Whole Blood (Level III) lists Pb data from two independent sources that exhibit poor between-method agreement. The original Pb concentration data as presented on the Certificate of Analysis are 641 µg L^{-1} ± 35 µg L^{-1} and 554 µg L^{-1} ± 22 µg L^{-1}. These data were combined using the Type B on Bias method to arrive at a mean and expanded uncertainty value (coverage factor, $k = 2$). The resultant data were then converted to units of mass fraction using a density of 1.056 g mL^{-1} to arrive at a value of 0.566 mg kg^{-1} ± 0.051 mg kg^{-1}. The certified Pb value and expanded uncertainty for SRM 966 Toxic Elements in Bovine Blood displayed in Fig. 5 has been corrected for density using a value of 1.0544 g mL^{-1}.

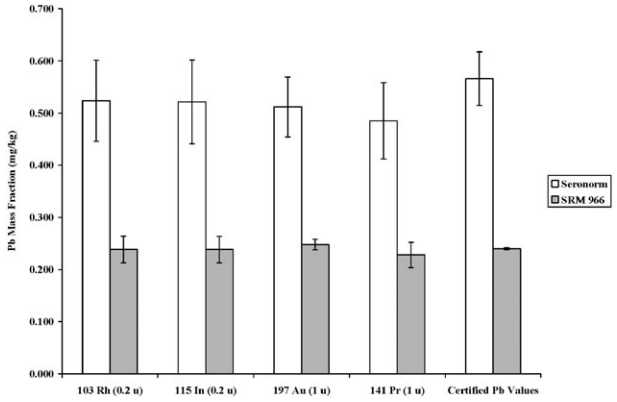

Fig. 5 Comparison of Pb mass fraction data (mg kg^{-1}) as a function of internal standard reference at 0.2 u mass resolution (^{103}Rh and ^{115}In internal standards) and at unit mass resolution (^{197}Au and ^{141}Pr) with certified Pb mass fraction data for CRM Seronorm Trace Elements in Whole Blood (Level III) and SRM 966 Toxic Elements in Bovine Blood. All Pb data was collected at a mass resolution of 0.2 u.

is different from that of the analyte because separate experiments would need to be performed measuring identical internal standard elements at appropriate concentrations under different resolution settings. Also, the Pb concentration uncertainties are large relative to the uncertainties for Pb content in both certified reference materials. We might speculate that the poorest analyte/internal standard data measurement conditions likely to foster a decrease in precision at the *X*-intercept in a standard additions experiment would probably be realized for an analyte measured at or near its detection limit which was referenced to an internal standard measured at high resolution, or alternatively when both measurements for an analyte/internal standard pair are collected at high resolution, where poor quadrupole peak shapes (relative to unit resolution mode) may be a controlling factor.

The data generated for Cd and Pb in SRM 966 were deemed encouraging because they were produced from a standard additions calibration curve generated using Seronorm whole blood. This suggests that blood samples from different physiological systems (and likely, moderately different matrices) can be predicted from the same standard additions calibration curve, with any subtle matrix effects being compensated for by the internal standards. This should allow for benchmarking the accuracy of whole blood measurements for numerous physiological systems (animals, humans, *etc.*) with current-issue whole blood CRMs. The work described here has implications for high throughput clinical experiments that require high-accuracy, matrix-matched calibration schemes because numerous unknown samples and control samples of similar matrix can be run against the slope of a single standard addition curve.

We have found that the experimental design and working curve range for this ratio-based standard additions calibration scheme is critical. An inherent problem was uncovered upon further analysis of our Cd data. The measured $^{111}Cd^{+}/^{89}Y^{+}$ ratio of the highest concentration spike in our standard additions calibration curve (not shown) was approximately 150× higher than the measured $^{111}Cd^{+}/^{89}Y^{+}$ ratio of the unspiked sample. This is not an ideal working range for implementation of the method of standard additions. We have observed that eqn. (3) cannot be applied in this situation when measuring unknown or control samples against the standard additions curve, because the resultant standard error of the *y* estimate (V_y) will inflate, *i.e.*, the prediction uncertainty will increase substantially. So, even though we were able to produce accurate Cd data with reasonable uncertainties from replicate analyses of three aliquots of material, we could not apply eqn. (3) to compute an uncertainty for each individual determination and later combine them. This effect was less pronounced for the Pb regressions because the spike factor was more reasonable (high concentration spike factor of approximately 7×) and not an issue for the SRM 1947 certification experiment because the experimental design placed the high concentration spike only about a factor of 2–20 higher than the corresponding unspiked sample (for example, refer to the calibration curves for Cu, As and Fe in Figs. 1–3, respectively).

To date we have successfully applied this calibration technique and analytical methodology to the determination of multiple elements in clinical samples without sacrificing accuracy and high throughput. This includes successful measurement of some of the more problematic quadrupole ICP-MS elements, like V, that benefit from reduction of the oxychloride interference $^{16}O^{35}Cl^{+}$, which can be effectively removed by utilization of both reactive gas chemistry and collision induced dissociation processes employing ammonia as the reagent gas in a helium buffer gas. It is beyond the scope of the present work, however, to present and interpret the multi-element data collected for specific projects: this work will be presented in a separate study. Here our purpose is to discuss validation of this relatively unique approach to standard additions calibrations using modern ICP-MS technology and demonstrate the accuracy of the ratio-based standard additions approach as applied to metrological and critical, high sample load analytical problems. Future work will focus on applying the ratio-based standard additions quantification scheme to SRM 1598a Inorganic Constituents in Animal Serum, which is currently in preparation.

Conclusions

A ratio-based, multi-element calibration scheme based on the method of standard additions and collision cell ICP-MS has been developed and applied to the certification of trace elements in SRM 1947, a candidate fish tissue Standard Reference Material. Concentrations for multiple trace elements in SRM 1946 Lake Superior Fish Tissue were successfully predicted using the slope of a standard additions calibration curve derived from SRM 1947 Lake Michigan Fish Tissue, demonstrating that the technique is robust and tolerant of slight variations in the sample matrix. Formulae for calculating regression and prediction errors based on quotient propagation of errors were put forward and caveats on their use with an improper standard additions spiking regimen were described. The scheme was also successfully applied to the quantification of Cd and Pb in clinical, whole blood CRM samples. The sample preparation and calibration scheme for the clinical samples proved to be robust and tolerant of samples from different physiological systems (bovine and human blood) and, correspondingly, slightly different matrices. The streamlining advantages of implementing collision-cell ICP-MS with multiple internal standards and multiple resolution settings were highlighted, through testing of interference prone elements (*e.g.*, As, Se, *etc.*) and analyte/internal standard pairs where the analyte signal and corresponding internal standard signal was measured at different mass resolution settings. Finally, the benefits of coupling a gravimetric, internal standard ratio-based calibration scheme with the method of standard additions were discussed in detail from a historical and practical perspective and these benefits include non-quantitative sample handling, reduced number of analytical samples, simple adaptation to high sample load analytical problems and compensation for instrumental drift.

Acknowledgements

The authors would like to acknowledge Stefan Leigh (NIST Statistical Engineering Division) and Professor Andrew Ruhkin (University of Maryland, Baltimore County) for their

assistance with developing the regression and prediction uncertainty formulae. Professor David Zellmer (California State University Fresno) is also acknowledged for engaging in fruitful discussions with the authors, and providing an insight into sources of calibration and prediction errors for the standard additions technique. Finally, Therese Butler (NIST Chemical Science and Technology Laboratory) is acknowledged for assistance with the preparation of the custom multi-element spike solutions used throughout the course of this work.

References

1 J. D. Fassett and P. Paulsen, *J. Anal. Chem.*, 1989, **61**, 643A–649A.
2 J. J. Thompson and R. S. Houk, *Appl. Spectrosc.*, 1987, **41**, 801–806.
3 H. Vanhaecke, H. Vanhoe, R. Dams and C. Vandecasteele, *Talanta*, 1992, **39**, 737–742.
4 D. R. Wiederin, R. E. Smyczek and R. S. Houk, *Anal. Chem.*, 1991, **63**, 1626–1631.
5 D. E. Nixon and T. P. Moyer, *Spectrochim. Acta, Part B*, 1996, **51**, 13–25.
6 A. S. Al-Ammar, *Spectrochim. Acta, Part B*, 2003, **58**, 1391–1401.
7 K. Cottingham, *Anal. Chem.*, 2004, **76**, 35A–38A.
8 D. W. Koppenaal, G. C. Eiden and C. J. Barinage, *J. Anal. At. Spectrom.*, 2004, **19**, 561–570.
9 S. J. Christopher, *Description and Results of the 2001 Interlaboratory Comparison Exercise for the Determination of Trace Elements in Marine Mammals: Report to Participants*, National Institute of Standards and Technology, Charleston, SC, 2001.
10 A. M. Featherstone, A. T. Townsend, G. A. Jacobson and G. M. Peterson, *Anal. Chim. Acta*, 2004, **512**, 319–327.
11 E. McCurdy, Agilent Technologies UK Limited, 2004, personal communication.
12 C. E. Sieniawska, R. Mensikov and H. T. Delves, *J. Anal. At. Spectrom.*, 1999, **14**, 109–112.
13 M. L. Salit and G. C. Turk, *Anal. Chem.*, 1998, **70**, 3184–3190.
14 A. L. Rukhin and M. G. Vangel, *J. Am. Stat. Assoc.*, 1998, **93**, 303–308.