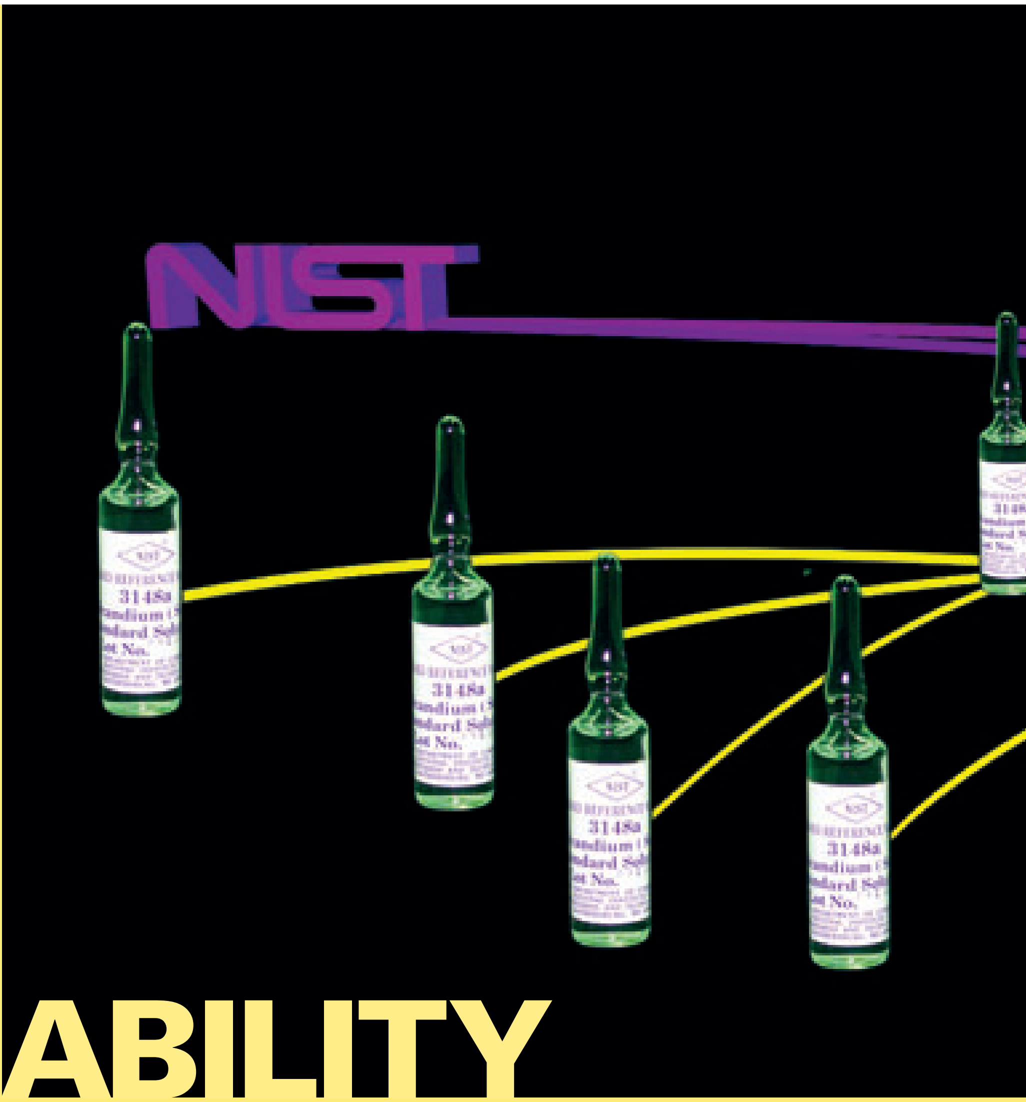

Elemental solution standards can be produced that are almost "clones" of NIST SRMs.

TRACEABILITY of Single-Element Calibration Solutions

Commercially available single-element solutions are used to calibrate the instruments with which millions of elemental determinations are performed each day. Atomic spectrometric techniques rely on calibration against these solution standards to determine the tens of elements in hundreds of samples in thousands of labs. The solutions are, or are prepared from, certified reference materials (CRMs), which are the basis for the accuracy of the measurement results. A certificate that asserts a claim of "traceability" usually accompanies these CRMs.

The claim of traceability aims to ensure that the certified value of the CRM is linked to appropriate national and international references, which are presumably stable and accurate and of higher metrological order. Properly established, this linkage couples the information content of the traceable value to the information content of the reference, and it can add immensely to the acceptance of the CRM. This linkage also connects the measurement results calibrated from the traceable value of a CRM to the larger network of measurements, which ultimately stems from the International System of Units (SI). The maintenance of the base and derived units of SI, together with traceability to SI, underpins contemporary metrology. SI provides a uniform, stable measurement scale, and traceability of measurement results to SI ensures their comparability over time and space.

To provide suitable references for the millions of atomic spectrometric determinations made daily, the National Institute of Standards and Technology (NIST) disseminates elemental solution standards through its standard reference materials (SRMs) program; SRMs are CRMs distributed by NIST. We will describe the concept of traceability and one approach by which traceability to these references can be achieved by directly comparing solutions against them. The approach we describe is based on work from our laboratory using a method we refer to as high-performance inductively coupled plasma-optical emission spectrometry (HP-ICP-OES; *1–4*).

Traceability

The concept of traceability predates our contemporary global metrology system. Originally, the king's foot was a measure of length; then, traceable artifacts the same length as his foot were disseminated to permit measurements to be made with ease, thus saving the king's foot a lot of wear and tear. The *International Vocabulary of Basic and General Terms in Metrology* defines traceability as the "property of the result of a measurement or the value

Marc L. Salit
Gregory C. Turk
National Institute of Standards and Technology

of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties" (5). This definition, together with agreed-upon units of measure, ensures that a traceable result (a value and its associated uncertainty) measured in the absence of bias can be compared with other traceable results, also measured in the absence of bias. Such results may arise from comparison to the value of the same stable artifact in different places at different times, or from comparison to the values of different artifacts, whose values are also traceable.

Comparability of the results is ensured by traceability. The ability to compare a result to other results, expectations, or specifications is essential. A measurement result that cannot be compared is of no value. Even if results are to be compared only within a given laboratory, those results are useful only by virtue of their values relative to other results.

The extent of space and time over which results can be compared depends on the scope of the reference to which they are linked or to which they are traceable. Local traceability ensures local comparability. A measurement result traceable to a reference held in a particular laboratory can be meaningfully compared only to other measurement results traceable to that reference. If that reference is itself traceable to an external reference (to which other references may also be traceable), then traceability to that external reference may be established. In this manner, measurement results can then be compared among laboratories. This network rises from standards held at individual laboratories through those held at national and international levels.

Comparability between national and regional measurement systems is being established in part through the 1999 Comité International des Poids et Mesures (International Committee for Weights and Measures; CIPM) multilateral mutual recognition arrangement. This agreement, signed by the 38 member nations of the Meter Convention, was a response to the need for an open, transparent, and comprehensive scheme to give users reliable quantitative information on the comparability of national metrology systems (www.bipm.fr/enus/8_Key_Comparisons/mra.html). Through this horizontal comparability between nations, the vertical traceability of a measurement to a national reference may link it to international comparability, which is potentially important for international trade, commerce, and regulatory affairs, as well as global science.

Of increasing relevance to chemical analysis is laboratory accreditation based on International Organization for Standardization Guide 17025 standards that require a clear demonstration of traceability. In the United States, government regulations or laboratory quality systems often require that the results of an analysis be "traceable to NIST," an ill-considered phrase that, more properly stated, means "traceable to the values on NIST measurements or standards." NIST policy on traceability explicitly states: "Other organizations are responsible for establishing the traceability of their own results or values to those of NIST or other stated references" (www.nist.gov/traceability). Additionally, the policy states that NIST "[a]sserts that providing support for a claim of traceability of the result of a measurement or value of a standard is the responsibility of the provider of that result or value, whether that provider is NIST or another organization, and that assessing the validity of such a claim is the responsibility of the user of that result or value." This caveat warns that the users of materials that claim traceability must be able to evaluate those claims, and thus it gives tremendous importance to the transparency of a traceability claim.

The analyst must recognize that traceability only requires two things: a suitable reference to compare to and a measurement method for the comparison. Of course, an essential part of a traceable result is its uncertainty, which must consist of the uncertainty of the value of the reference combined with the uncertainty of the comparison. An assessment of traceability evaluates the validity and suitability of the reference, the appropriateness and reliability of the comparison, and whether the uncertainty of the traceable value is valid.

Elemental solution standards

In the United States, designating a calibration solution for elemental analysis as traceable usually means that the certified value of that solution is traceable to the certified value of the matching elemental solution from the NIST 3100 series of SRMs. These 69 single-element solution SRMs serve as national standards for elemental solution mass fraction. These SRMs are prepared in bulk and are batch-certified for the mass fraction of the constituent element (typically 10 mg/g). The target range for the relative uncertainty in the certified mass fraction for these SRMs is ±0.3%.

As part of our process to certify the 3100 series, candidate SRM materials are compared against well-characterized primary materials. Through this comparison, traceability of the SRM values to the SI is established. HP-ICP-OES has been our comparison method of choice for this process since 1997. This method uses unmodified, readily available, mature, reliable commercial instrumentation; it has been demonstrated to have very small comparison uncertainty, ruggedness with respect to bias (small changes in conditions do not bias the results), economy, sufficient linear dynamic range, and a high degree of inter-element selectivity. The HP-ICP-OES method has been applied to the analysis of single-element solutions for 64 elements (*3*), the certification of the major constituents of a high-temperature alloy (*4*), and the characterization of an $LiAlO_2$ ceramic material (*2*). The method was used in an international key comparison (K8: Monoelemental Calibration Solutions) of the Comité Consultatif pour la Quantité de Matière (Consultative Committee for the Amount of Substance) of the CIPM, with excellent results that validated it against a number of other analytical techniques (*6*). A similar approach using ICPMS for the multielement analysis of environmental CRMs has been reported (*7*).

Pairing of the 3100 series with a rugged and practical comparison method enables the manufacture of traceable calibration solutions. This traceability linkage couples the information content of the SRM to the traceable solution. Thus, a manufacturer of a CRM could leverage this approach, relying on comparison against the SRM, and avoid duplication of the work done at NIST. The 3100 series solutions are themselves traceable to the mole, the SI unit for chemistry. The well-characterized primary materials, against which the SRMs are compared, are realizations of the mole for the elements. Comparison against the SRM can yield a value for a CRM, also now traceable to the mole. Production of a rhodium solution CRM is a good example. Rhodium metal is extremely difficult to dissolve, and soluble salts of rhodium are difficult to assay with small uncertainty. A commercial or in-house CRM solution could be prepared from a soluble salt, and a traceable value derived by comparison against SRM 3144, the rhodium standard solution SRM (*8*, *9*).

Unquestionably, the traceable mass fraction of analyte in the CRM could not have smaller uncertainty than the SRM. The 3100 series solutions are optimized to have stable mass fractions with small uncertainties. However, the quality of a calibration solution is derived from a number of parameters, not only the uncertainty of the mass fraction. Solution purity, stability, packaging, and concomitant anions all play a role in fitness for purpose, or quality. A traceable CRM could be of either higher or lower quality than the SRM.

For example, many of the 3100 series solutions are packaged in glass ampoules that enhance stability by limiting transpiration but could induce contamination. A CRM could have a certified value traceable to the SRM and have higher purity, making it more suitable for multielement analysis. Or, a solution CRM might be prepared from a poor grade of material containing many impurities. The traceable certified value might be excellent—accurate, with small uncertainty—yet the material may be inappropriate for its application and of low quality.

How to do it

HP-ICP-OES exploits two strengths of array-detector-based spectrometers: simultaneous measurement of the analyte and internal standard, and time-correlated off-line background correction (*3*). Moderately high analyte mass fractions and signal levels provide practical immunity from blank and spectral interferences, and a drift-correction approach mitigates low-frequency noise. These strategies typically yield results with variability <0.1%, which is on the order of that associated with sample handling. Sample-handling variability is quantified with an experiment that includes the analysis of replicate preparations. The HP-ICP-OES method, or one like it using a different measurement tool for comparison, might be used by those wishing to prepare solutions traceable to NIST SRMs—in particular, by commercial CRM manufacturers whose products are designed to ensure the accuracy and comparability of most elemental analysis results.

We have developed a spreadsheet tool that leads the analyst through the HP-ICP-OES comparison to establish traceability of the value of a CRM to the value of a NIST 3100 series SRM. The essential elements of this comparison include a fixed experimental design; a method development experiment that establishes the sensitivities for the analyte and an internal standard; a gravimetric sample preparation scheme; a fixed set of emission intensity measurements of both elements; and the calculation of the traceable value and its uncertainty. The spreadsheet also includes useful diagnostic charts that permit visualization of the instrument drift and the noise correlation between the analyte and internal standard. The spreadsheet tool, an example data set, and instructions are available in Supporting Information.

A measurement result that cannot be compared to other results, expectations, or specifications is of no value.

Four separate preparations are made of the desired element from the 3100 series SRM and the single-element solution to be compared (the test sample). The 4 preparations of the SRM are designed to consume ~10 g of solution (many of the 3100 series solutions are deployed in packages of 5 single-use 10-g glass ampoules). Ideally, the four preparations of the test sample are selected from different portions of the batch, so the population of results from the test sample will include a measure of any heterogeneity. This replication permits evaluation of several uncertainty components associated with the preparation of the solutions and acts to validate the method and its results. The analyst must take proper care to develop the instrumental method (wavelengths, spectral-background-correction approach and parameters, and integration times) such that the measured emission intensities have high S/N and selectivity. A useful reference for method development is the study reporting HP-ICP-OES analyses for 64 different elements (*3*).

Sample prep. Analyte and internal standard mass fractions are selected to yield an emission intensity ratio near unity. For some ICP-OES instruments, this facilitates simultaneous integration, in which time is used to extend the dynamic range of a spectrometer, and limits the impact of deviation from linear intensity response.

Because the analyte and internal-standard spectral lines are likely to have different sensitivities, a preliminary experiment is performed to roughly determine the sensitivity ratio. Solutions of the analyte and internal standard, whose mass fractions are approximately known (10% uncertainty is adequate), are measured under the selected operating conditions, and the emission intensities are recorded. These solutions should be well within reasonable operating levels for the instrument; for most elements, 10 µg/g is appropriate. Once the sensitivity ratio is established, a mass ratio (mass of analyte per mass of internal standard) that yields unity intensity ratio can be calculated. Then, an internal standard spike solution of appropriate mass fraction can be prepared.

Once the spike is added to the analyte, the mass ratio is established. This is the only sample preparation data value used in calculating the result, and it is unaffected by any subsequent dilutions. Typically, the spiked samples are diluted when introduced to the plasma, in order to yield an appropriate signal level that is comfortably within the working linear response range of the instrument and sufficiently larger than any anticipated blank or interference level. This dilution can be performed by eye using a digital pipette, and it is not necessary to record any values associated with this dilution.

All masses whose uncertainties propagate to the results are >2.5 g. This amount permits the use of common top-loading balances, which typically have 1.5-mg-repeatability standard deviations. This keeps the relative uncertainty of weighing to <0.1%, which is on the order of our ability to prepare duplicate preparations. Aliquots of ~2.5 g are used for both the SRM and the test samples. The scheme assumes that the nominal mass fractions of the SRM and the test samples are within an order of magnitude of each other. Aliquot masses are adjusted to deliver the same mass of analyte from each solution.

An internal standard solution is prepared so that the spike masses are ~20 g. The mass fraction of this solution need not be accurately known, as only the relative masses of spike weighed into the reference and test solutions appear in the calculation of results. The mass used need not be precisely 20 g, and the solution is typically delivered with a graduated cylinder. The 20 g of spike solution is easily weighed with negligible uncertainty and typically dilutes the high-mass-fraction (10-mg/g) SRM. When two solutions with high mass fractions are mixed, a precipitate might form; dilution helps mitigate this. Final dilution from the spiked sample is also facilitated by this initial dilution. Moreover, the lower the mass fraction of the spike, the less the relative loss (from missed droplets, evaporation, or aerosolization) during solution handling.

Intensity measurements. Intensity measurements are performed sequentially on the eight samples, alternating the SRM and test samples (the run order is SRM sample 1, test sample 1, SRM sample 2, test sample 2, etc.). This sequence of measurements is repeated five times. Each is typically the mean of several replicate intensity measurements. The short-term noise of the intensities, expressed as the relative standard deviation of the replicate measurements, is a good diagnostic measure of the sample input performance. The short-term noise of the ratios, and its improvement over the intensity noise, helps to differentiate among the sources of measurement noise; the correlation of an-

Table 1. Sensitivity test.

	Sensitivity solution mass fraction (μg/g)	Measured emission intensity	Sensitivity $[(\mu g/g)^{-1}]$	Target working mass fraction (μg/g)	Sensitivity ratio
Cobalt	10	1,670,869	167,086.9	5.98491	0.037901
Scandium	5	22,042,376	4,408,475	0.226836	

alyte and internal standard signals is an indicator of performance and sets the expectation for uncertainty.

Calculations. The straightforward calculations are based on Equation 1 in Reference 6. Intensity ratios are calculated for each repeat measure of each solution; these ratios are drift-corrected and summarized for each sample. Mass ratios are calculated from the preparation data, and the ratio of the intensity ratio to the mass ratio is calculated for each sample. A single calibration factor is determined from the summary of the SRM measurements and is used to calculate the traceable mass fraction of the test sample.

The uncertainty budget includes components for dispersion of measurements of the SRM and the test sample (due to variability of the intensity ratio measurements and preparation and weighing) and for the uncertainty of the certified value of the SRM. A coverage factor is selected to expand the standard uncertainty to a 95% confidence interval.

Test experiment

Two ampoules of cobalt SRM 3113 (lot 000630, certified mass fraction 9.996 ± 0.024 mg/g) solution were selected to demonstrate the approach. One of these ampoules was diluted by ~0.5% (45.76 mg of 2% volume fraction HNO_3 was added to 10.30459 g of SRM solution) to prepare the synthetic test sample. The test sample mass fraction was 9.952 ± 0.0043 mg/g. Cobalt mass fraction was determined by HP-ICP-OES comparison of the diluted sample against the unadulterated SRM. Details of the experiment can be found elsewhere (*3*). Table 1 reports the results of the sensitivity test: the mass fractions of the sensitivity solutions, measured intensities, sensitivities, working mass fractions (nominal mass fractions of analyte and internal standard to be introduced to the plasma), and sensitivity ratio. The working mass fractions were calculated to yield an operating signal level of 10^6 counts/s.

The target and actual weights for the eight solutions prepared can be found in Supporting Information. The target weights are based upon the sensitivity test, an SRM aliquot of 2.5 g, and a minimum spike weight of 20 g. Target weights are nominal values. Variations of as much as tens of percent are well within expected norms when the solutions are prepared. Typical variability is observed in these data. Approximately 60 g of solution were prepared for analysis. Based upon the working mass fraction, the nominal weight of spiked solution to dilute to 60 g is calculated as 0.32 g. The 5 replicate measurements of the 8 samples took place in <2 h. Five repeat integrations of 8 s each were measured for each replicate measurement of each sample. Total photon integration time was 26 min 40 s. At every sample change (40 of them), 90 s elapsed to permit the new solution to traverse to and equilibrate in the spray chamber and reach the plasma. The analyte and internal standard signals were highly correlated. A sevenfold reduction in the noise in the ratio compared to the analyte signal can be observed in Figure 1.

A straight-line calibration model is fitted and forced through zero for the four SRM solutions. The slope of this line is estimated as the mean of the slopes calculated for each of the four solutions, resulting in a single calibration factor. Mass fractions for all solutions (treating each SRM solution as an unknown) are then calculated from this calibration factor; Figure 2 presents results before and after drift-correction and reference lines for the mean mass fraction and uncertainty of the SRM. The center of the gray reference bar is the reference value for our synthetic test sample, and the thickness is its uncertainty. The uncertainty of the SRM is excluded from all results graphed for the test sample to allow only the comparison to be assessed. The dashed black lines represent the mean and standard uncertainty of the HP-ICP-OES results (9.950 ± 0.0022 mg/g before drift-correction and 9.950 ± 0.0026 mg/g thereafter).

Several observations can be made directly. When the data for the individual solutions are drift-corrected, the dispersion is smaller. However, the dispersion among the results of the four solutions (for either the SRM or the test sample) is essentially unaffected by drift-correction. The dispersion among the four preparations of a solution dominates the dispersion of the measurements. This preparation effect has been documented and arises from uncertainty in solution handling and weighing, limitations of the balance, and evaporation of the solution while it is

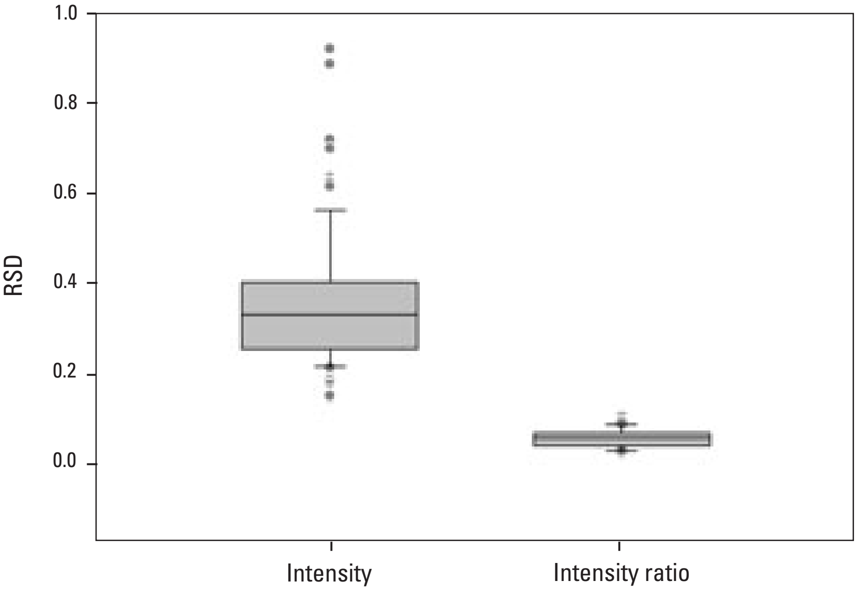

FIGURE 1. S/N improvement expressed as relative standard deviation with time-correlated internal standardization.

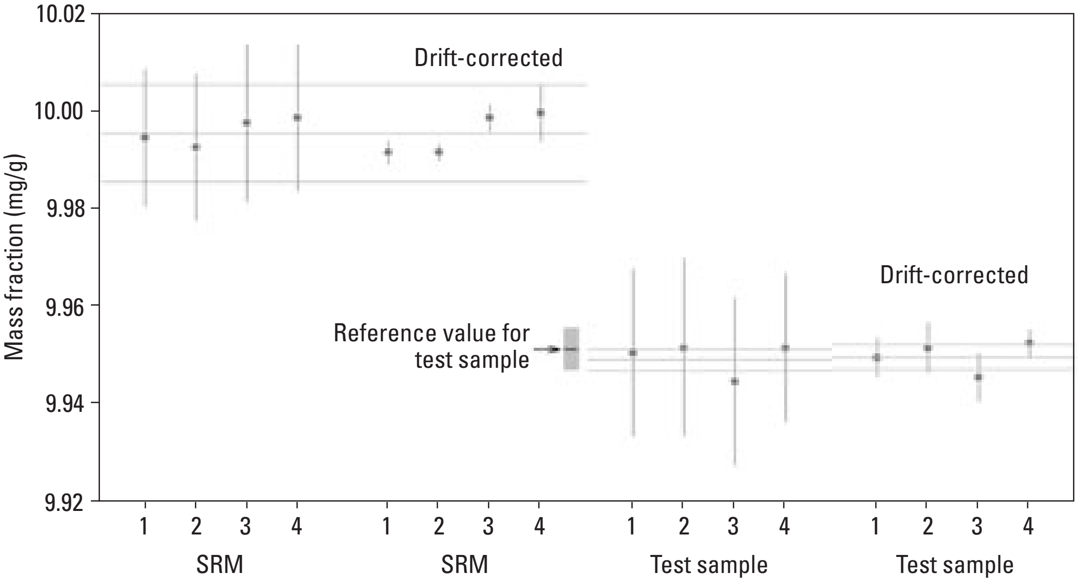

FIGURE 2. Results of test–experiment comparison.

being weighed. The magnitude of this effect is ~0.03% (*3*). The result for the synthetic test sample from the comparison is consistent with the reference value.

The uncertainty budget before and after drift-correction for the comparison is presented in Table 2. This simple yet rigorous budget includes components of uncertainty for the measurement of the test sample and SRM 3113. These standard uncertainties are evaluated as the standard deviation of the mean of the results from the four preparations. The variability of the intensity ratios is convolved into the dispersion of the replicate preparations, as is the variability of preparation. This "method-validation" approach to uncertainty evaluation is a useful feature of the experiment, bundling together the evaluation of individual uncertainty components into evaluation via replication.

The dominant uncertainty in the traceable value calculated from the comparison arises from the uncertainty of the certified value of the SRM. This value is ~4× larger than the uncertainty due to the comparison. It is a general rule of thumb that uncertainty components 3× smaller than the largest components can be regarded as insignificant (*11*). This rule of thumb applies here: We observe an insignificant difference between the uncertainty of the certified value of the SRM and the uncertainty of the result for the test solution. The results of the test experiment demonstrate that establishing traceability can be done without inflating uncertainty and in less than half a workday. HP-ICP-OES is an effective tool to link the value of a test solution, commercial CRM, or in-house standard to global comparability. When traceable materials are used to calibrate a valid method, the results of analyses using that method can be compared with other results, thereby giving them real value.

Table 2. Results and uncertainty budget.

	Before drift-correction		After drift-correction	
	mg/g	Degrees of freedom	mg/g	Degrees of freedom
Observed value of test solution	9.9496		9.9501	
Uncertainty sources				
SRM 3113	0.0014	3	0.0021	3
Test sample	0.0017	3	0.0016	3
Combined uncertainty of measurements only	0.0022		0.0026	
Uncertainty of certified value of SRM 3113	0.0100	8	0.0100	8
Combined standard uncertainty	0.0102	9	0.0103	9
Coverage factor	2.26		2.26	
Expanded uncertainty	0.0230		0.0232	

Marc L. Salit is a research chemist and Gregory C. Turk heads the spectrochemical methods group at the National Institute of Standards and Technology. Address correspondence about this article to Turk at gregory.turk@nist.gov.

References

(1) Salit, M. L.; Turk, G. C. *Anal. Chem.* **1998**, *70*, 3184–3190.
(2) Salit, M. L.; Vocke, R. D., Jr.; Kelly, W. R. *Anal. Chem.* **2000**, *72*, 3504–3511.
(3) Salit, M. L.; et al. *Anal. Chem.* **2001**, *73*, 4821–4829.
(4) National Institute of Standards and Technology. SRM 1775/2175 Certificate of Analysis. NIST: Gaithersburg, MD.
(5) International Organization for Standardization. *International Vocabulary of Basic and General Terms in Metrology*, Section 6.10, ISO: Geneva, Switzerland, 1993.
(6) *Metrologia* **2002**, *39*, Tech. Suppl. 08002.
(7) Turk, G. C.; et al. *Fresenius' J. Anal. Chem.* **2001**, *370*, 259–263.
(8) Beck, C. M., II; et al. *Anal. Chem.* **1993**, *65*, 2899–2902.
(9) Gordon, C. L.; Schlecht, W. G.; Wichers, E. *J. Res. Natl. Bur. Stand. (U.S.)* **1944**, *33*, 363–381.
(10) Taylor, B. N.; Kuyatt, C. E. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*; NIST Technical Note 1297, 2nd ed.; National Institute of Standards and Technology: Gaithersburg, MD, 1994; http://physics.nist.gov/cuu/Uncertainty/index.html.
(11) Eurachem/CITAC Guide: *Quantifying Uncertainty in Analytical Measurement*, 2nd ed.; Laboratory of the Government Chemist: London, 2001.