Elemental Analysis Manual

for Food and Related Products

The following is a section of the Elemental Analysis Manual for Food and Related Products.

For additional information and to view other sections of the manual, visit the Elemental Analysis Manual for Food and Related Products web page at http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm2006954.htm.

U.S. Department of Health & Human Services

FDA U.S. Food and Drug Administration

Elemental Analysis Manual

for Food and Related Products

3.5 Reference Materials

September, 2014

Authors: William C. Cunningham
Stephen G. Capar

Table of Contents

Reference materials (RMs) are used for method validation, verification of correct use of a method, calibration, and quality control. The sections below provide information on some of FDA's elemental analysis thinking on the use of RMs but additional information is available[1-3].

Reference Material (RM)—Material or substance one or more of whose property values are sufficiently homogeneous, stable, and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials[4].

Certified Reference Material (CRM)—Reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes its traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence[4].

In-house Reference Material (in-house RM)—Reference material developed by a laboratory for its own internal use.

CRMs and in-house RMs are simply types of RMs. A CRM is a RM with an associated certificate that satisfies traceability requirements. When a certificate expires, the material in the unit (or container) continues to be an RM but is, by definition, no longer traceable and therefore no longer a CRM. The name "standard reference material" (SRM) is sometimes used generically when discussing any CRM but is actually a trade name that the U. S. National Institute of Standards and Technology (NIST) uses for a CRM.

3.5.1 REFERENCE MATERIAL USE FOR QUALITY CONTROL

RMs are analyzed with a batch of samples to verify the accuracy and overall performance of the analysis. CRMs provide traceability. Other non-certified RMs, including expired CRMs and RMs with established consensus values (*e.g.*, Roelandts and Gladney[5]), can be used to demonstrate repeatability. CRMs may be used in any RM application but must be included, alone or in combination with non-certified RMs, in regulatory analyses. CRMs are typically not required for investigational and surveillance analyses.

Instructions are given on RM certificates for the recommended minimum analytical portion mass, the procedure for determining mass at a reference moisture condition, storage requirements, etc. The use of z-scores (see §3.5.3) is an accepted procedure for demonstrating the quality of results.

RMs are chosen to closely match the matrix and analyte concentration of interest. However, this is often difficult because of the relatively small variety of RMs available. For FDA labs, the difficulty in choosing an appropriate RM is compounded by the large variety of food matrices and numerous analytes.

The following sections provide guidance in obtaining or preparing RMs.

3.5.2 IN-HOUSE REFERENCE MATERIAL DEVELOPMENT

An in-house RM is usually developed when the matrices or reference levels of commercially available RMs do not closely match the samples to be analyzed or when an ample supply is desired that will be available for many years. The steps for development of an in-house RM for elemental analysis are outlined below.

3.5.2.1 SELECTION

Ideally, the material will be available in an ample supply, with minimal cost, and needing little or no preparation. Materials requiring freezing, pulverizing, sieving, blending, sterilization,

packaging, etc., should be selected only if necessary.

Logically, the material would be expected to be homogeneous and stable with a long shelf life. If refrigerated storage is required, cryogenic temperatures (-80 C or colder) is assumed. Analyte levels, interference issues, and analysis difficulty need to suit the purpose of the RM. Challenging materials having interferences or digestion complications are generally undesirable but are useful to demonstrate ruggedness.

3.5.2.2 ANALYTICAL

Analysis data are used to characterize the material's behavior, set analyte levels, determine minimum analytical portion mass, and set uncertainties.

A procedure must be defined, for bringing the material to a reproducible moisture condition so the correct analytical portion mass can be known. Typical moisture conditions include freeze-dried, oven-dried, as received, desiccator-dried, equilibrium mass state, and reconstituted.

Analyte levels are best established using a variety of analytical techniques and methods in different laboratories and by different analysts. Use of multiple sets of data in this way compensates for the small and unknown amounts of method bias that are always assumed to exist.[6-7] Limits can be deduced on how large the bias may be and the certificate uncertainties can be adjusted accordingly.

For each analyte, the reference value should be established using a combination of at least 2 different analytical techniques or laboratories. The use of a consensus mean approach such as by DerSimonian and Laird[8] is recommended. FDA has used this and other approaches commonly used in CRM production at NIST. [9]

When analysis data are available for only one source (method/laboratory/analyst), then there is no experimental confirmation for the reference values and it is much more difficult to arrive at a reasonable uncertainty component for bias.

3.5.2.3 RANDOM ERROR AND HOMOGENEITY

A study specifically to address nonhomogeneity would not, necessarily, be necessary to set reference values for the analytes of interest because it will be reflected in the observed random error. However, it is quite useful when evaluating reported data, setting the minimum analytical portion mass specification, and assigning reference values and uncertainties.

Analyte nonhomogeneity can be defined as the RSD from analyte variations within the material (RSD_{nonhom}) and expressed as relative percent. It is a characteristic that, as shown in equation 1, combines with random error from the measurement process (RSD_{meas}) to give the total, or observed, RSD (RSD_{obs}). [10]

$$\left(\mathrm{RSD}_{meas}\right)^2+\left(\mathrm{RSD}_{\mathrm{nonhom}}\right)^2=\left(RSD_{obs}\right)^2 \qquad \text{3.5 Equation 1}$$

Ideally, nonhomogeneity will be negligible for all analytes and RSD_{obs} will be due entirely to analytical factors. It should be expected, though, that nonhomogeneity will be significant for at least some analytes and be an increasing issue as the analytical portion mass decreases. When developing an in-house RM, the objective is therefore to find the smallest mass that will result in

a trivial amount of nonhomogeneity for as few analytes as is practical. The minimum mass specified for commercial RMs is often ~300 mg.

Use of equation 1 is straightforward when RSD_{meas} is known. It enables several pieces of information to be discovered about the RM material, especially when measurement results are available from different sources (different methods, laboratories, analysts).

If RSD_{obs} is greater than RSD_{meas}, then nonhomogeneity can be calculated by re-arranging Equation 1. RSD_{nonhom} is a characteristic of the material whereas RSD_{meas} is a characteristic of a laboratory/method. This means that RSD_{nonhom} should be the same for all laboratories and is logically no larger than the smallest RSD_{obs}. Differences in RSD_{obs} are also therefore due to variations in RSD_{meas} at the different laboaratories.

If RSD_{obs} is less than the laboratory's reported uncertainty, the laboratory has overestimated their uncertainty. This is not typically considered a serious problem because all laboratories wish their performance to be at least as good as they claim so it is common for uncertainties to be (conservatively) overestimated.

Note - the following procedure is used for CFSAN in-house reference materials to decide how much larger RSD_{obs} must be relative to RSD_{meas} to conclude RSD_{nonhom} has been detected:

The adequacy of random error alone to account for an observed data distribution can be evaluated on the basis of the integral of the distribution function $P_x(\chi^2,v)$ from $x^2 = \chi^2_{(observed)}$ to $x^2 = \infty$, where χ^2 is chi-square distribution and v is the number of degrees of freedom[11]. Here, the integral of the distribution function is referred to simply as probability. When the probability is ≤10%, a nonhomogeneity component can be calculated.

In general, nonhomogeneities that are equal to the associated random measurement uncertainties would be expected to have probabilities <10%. Therefore, when a probability is >10%, the nonhomogeneity is known to be less than the random measurement uncertainty and the latter can be taken as an upper limit for nonhomogeneity. However, when RSD_{obs} is lower than the random measurement uncertainty, RSD_{obs} is taken instead as an upper limit for nonhomogeneity.

3.5.2.4 UNCERTAINTIES

Uncertainties need to be included with the analyte reference values. Very commonly, uncertainties are reported at a 95% confidence level, which is sometimes referred to as a "two-sigma" uncertainty. Methods for determining uncertainties vary and depend on the data sets and associated analytical information.

Caution is necessary when setting uncertainties and especially when using generic ("canned") routines such as noted above (section 3.5.2.2). Since the uncertainties given on the certificate will be used to judge individual analysis results, the RM uncertainties must be appropriate for individual test portions and not just for the bulk material. The concern here is associated with nonhmogeneity. If a large amount of data are available from many sources, analyte levels in the bulk RM material can be known extremely well (i.e., known with very small uncertainty). It is assumed, however, that RM uncertainties are used for routine quality assurance to confirm

individual analysis results. Therefore, the certificate uncertainties should instead predict the errors expected for individual analyses.

The significance of nonhomogeneity in this regard stems from Equation 1, the fact that nonhomogeneity will affect the analysis results, and that nonhomogeneity is a characteristic of the RM and not of the analysis. These put responsibility on the RM producer to propagate nonhomogeneity in the certificate's uncertainty or provide it separately so an analyst may combine nonhomogeneity with the RM uncertainty before evaluating their analysis results. FDA propagates nonhomogeneity in the elemental analysis in-house RM uncertainties.

In a practical sense, a useful rule when assigning an RM uncertainty is that it should probably not be any smaller than the smallest standard deviation associated with the data from which a reference value is derived. This follows from the likelihood that accurate data are used to generate the reference values. This, of course, would not always be true, with an obvious exception being for RM solutions.

The final caution discussed here is that as uncertainty decreases, the assignment of appropriate uncertainties becomes increasingly important because analyses will need to be increasingly accurate. This is quantified below (section 3.5.3 Equation 2).

3.5.2.5 INSTRUCTIONS

The certificate must contain instructions[12]. At a minimum, it will explain how to use the in-house RM, storage requirements, procedure for determining the reference moisture condition, minimum analytical portion mass, and reference values for the analytes.

3.5.3 REFERENCE MATERIAL RE-VERIFICATION

CFSAN has not assigned expiration dates for in-house RMs. Continued use of the RMs is subject to re-verification. This is the process that shows an RM is still fit for purpose and is based on observations and analytical results. The analytical results may be obtained specifically for re-verifying the RM or generated during its routine analysis. The re-verification may also be issued as a general update to the bulk material or may be limited to one or more analytes and/or be applicable only within one laboratory.

Whereas CRMs are used to demonstrate traceability, in-house RMs demonstrate repeatability. As such, when measurement results agree with established levels, not only does this show the analytical process is correct but also that the analyte level(s) in the in-house RM are unchanged. Thus, the levels for analytes of interest are effectively re-verified with each use of the RM.

Re-verification is accomplished by showing that the measured results agree with the reference values, relative to the measurement uncertainties. Measurement uncertainties are ideally determined along with analyses but they could also be generically assigned for well-defined methods (e.g., 10% for some elements in EAM methods). They may also be set to zero, but this is conservatively restrictive.

Continuous monitoring of RM results is useful because changes in element levels are observed in a timely fashion. Plotting element levels over time will show trends. Visual inspection can be used to verify the absence of obvious evidence that would cause one to question a RM unit’s physical integrity. For example, change in color, presence of mold or seeing liquid when the material should be dry would disqualify an RM unit.

For re-verification, at least 2 analytical portions of the RM are analyzed concurrently with at least one analytical portion of a CRM. The CRM and RM results are compared with the certified and reference values, respectively, by using z-scores[13]. A z-score is equal to the difference between the result and certified value divided by the square root of the sum of the squares of the uncertainties from both the reference and analysis results (see Explanatory Note below). Absolute values are used for z-scores and interpreted as follows:

|z-score| of 2 or less is acceptable (agreement with reference value)

|z-score| between 2 and 3 is questionable (questionable agreement with reference value)

|z-score| of 3 or more is unacceptable (disagreement with reference value)

Re-verification is successful for an element if at least two-thirds of the z-scores are in the acceptable range and none are in the unacceptable range. Thus, when only 1 or 2 analytical portions of a CRM or RM are analyzed, every z-score must be in the acceptable range. When 3 analytical portions are analyzed, at least 2 of the z-scores must be in the acceptable range and one may be in the questionable range. For 4 or 5 portions, only one can be in the questionable range, etc.

If re-verification is unsuccessful, all analysis information is examined. Mistakes such as data entry or calculation errors may only need correcting. Analytical problems may require repairing equipment or obtaining new reagents. Unexplained findings may require reanalysis. Repeated failure to re-verify may indicate a faulty RM or CRM unit.

Explanatory note about z-scores with an example:

A z-score[13] indicates how many standard deviations a result is from the reference value. Use of a z-score to examine data quality is a standardized way to evaluate results and provides an additional perspective besides that given by recoveries, which do not account for the reference or measurement uncertainties, or precision, which accounts for only random error. For this application, the z-score is defined as:

$$Z = \frac{x_m - x_c}{\sigma} \qquad \textit{3.5 Equation 2}$$

where: $\sigma = \sqrt{\sigma_m^2 + \sigma_c^2}$

x_m = *measured analyte level*

x_c = *is the certificate (reference) level*

σ_m = *total combined measurement uncertainty (one sigma, corresponding to a confidence level of approximately 67%)*

σ_c = *certificate uncertainty (one sigma).*

(Example) The CRM reference value for an element is 45.7 ± 8.3 mg/kg and the CRM's certificate states the uncertainty is at a 95% confidence level. The results of analysis of three analytical portions are 41.6, 33.4, and 51.1 mg/kg. The method is well-defined, LOQ for this element is 5.0 mg/kg, and the decision has been made to assign 10%

uncertainty (at the one sigma, or 67%, confidence) for values above LOQ.

$$\sigma_c = 8.3 \div 2 = 4.15 \text{ mg/kg}$$

$$\sigma_m = 41.6 \times 10\% = 4.16 \text{ mg/kg}$$

$$\sigma = \sqrt{(4.15)^2 + (4.16)^2} = \sqrt{34.53} = 5.876 \text{ mg/kg}$$

$$z = \frac{x_m - x_c}{\sigma} = \frac{41.6 - 45.7}{5.876} = \frac{-4.1}{5.876} = -0.7$$

This calculation is performed for all three results.

z-scores: *For 41.6 mg/kg* = *0.7 (acceptable)*

For 33.4 mg/kg = *2.3 (questionable)*

For 51.1 mg/kg = *0.8 (acceptable)*

Two out of three are acceptable and none are unacceptable so the RM is re-verified.

3.5.4 REFERENCE MATERIAL SOURCES

(1) National Institute of Standards and Technology (NIST)

(2) National Research Council of Canada (NRC)

(3) European Reference Materials (ERM)

Collaboration of three European reference material producers:

a) Institute for Reference Materials and Measurements (IRMM) of the European Commission's Directorate General Joint Research Centre, Belgium

b) Bundesanstalt für Materialforschung und -prüfung (BAM), Germany

c) LGC, United Kingdom

(4) National Institute for Environmental Studies (NIES)

(5) International Atomic Energy Agency (IAEA)

(6) Resource Technology Corporation (RTC)

RM distributor (not a producer)

(7) COMAR

Internet location with a searchable database of CRMs (not a producer)

(8) Virtual Institute for Reference Materials (VIRM)

Internet location for CRM stakeholder interaction. Has searchable RM database. Not a producer.

3.5.5 IN-HOUSE REFERENCE MATERIAL CERTIFICATES

3.5.5.1 FDA COCOA POWDER (CP)

Current certificate of analysis: FDA Cocoa Powder Certificate (2012, amended 2013) Excel 2010 (.xlsx) file

FDA CP Certificate (2012) was issued May 16, 2012. With this issuance, FDA CP was revalidated for continued use as an in-house RM for quality control/quality assurance for determination of element mass fractions in food and other biological materials.

Development work for FDA CP started in 1994 when approximately 10 kg of commercially-produced cocoa powder from a single production lot was obtained from a local grocery store. The material was re-packaged into pre-washed amber glass bottles and has been stored at FDA CFSAN. Bottles are sent to FDA labs on request. Each bottle contains about 300 grams of cocoa powder.

Reference values are derived from analytical results that have been reported. Since data have been received on an on-going basis, the reference values have evolved over time. The first FDA CP certificate was issued August 9, 1996. Minor updates occurred in 1998, 1999, and 2000 and a major update occurred May 4, 2006. As expected, changes in numerical values have been quite small.

Several enhancements were realized with the 2012 certificate:

1. The certificate is presented electronically in the form of an Excel 2010 workbook named "FDA Cocoa Powder Certificate (2012).xlsx". It replaces the traditional hard-copy certificate but one may be obtained by printing the first three worksheets. The workbook is password-protected with changes permitted only in specific cells designed for user input. Calculations are therefore protected from being altered inadvertently.
2. Basis mass is relative to the material exposed to air at 30% relative humidity (RH), which is expected to be the median condition for typical FDA laboratories. An analytical portion is exposed to laboratory air for at least 2 hours before measuring the mass. If a mass uncertainty of ±1% is acceptable, then no further calculation is needed. If ±0.5% is desired, an adjustment is performed using the actual laboratory RH. This calculation can be done automatically using a tool within the certificate.
3. The data set used to derive the reference values has expanded considerably. Data are now available for 43 elements from 17 sources (methods, laboratories, techniques, etc.). Detection limits are provided for 6 additional elements.
4. Reference values were set using state-of-the-art "concensification" software developed at NIST for generating SRM certificate values. Values are provided for Confirmed Values (those confirmed via multiple sources) and Unconfirmed Values (those from a single source).
5. Reference values are provided in duplicate listings. One listing applies for the bulk material and is useful for evaluating means obtained by averaging results from replicate analyses. The second listing applies for individual analytical portions and is useful for

evaluating individual analysis results. The former is typical for commercial CRM certificates while the latter includes sample-to-sample non-homogeneity effects.

6. Users select the coverage factor for displaying uncertainties. For example, by entering the number one (1), the reference values will have standard uncertainties (1-sigma or about 67% confidence); or, by entering the number two (2), the reference values will have uncertainties at about 95% confidence, which is typical for commercial RM certificates.
7. The certificate has a z-score evaluation capability. When a user enters analysis uncertainty (e.g., 10%), ranges are automatically generated to show the acceptable and questionable ranges for results. And, a hard-copy z-score evaluation may be generated to include with a report of analysis, a user needs only to enter the individual results for an analysis.

REFERENCES

(1) European co-operation for Accreditation, Eurolab and Eurachem (EEE) Working Group on Reference Materials (2003) The Selection and Use of Reference Materials. EA 4/14 (rev00). Accessed September, 2014.

(2) Lawn, R., Roper, P., Holcombe, G., and Stuart, B. (2001) Application Notes for the Production of Low-Cost Quality Control Matrix Reference Materials. National Measurement System - Chemical and Biological Metrology. Accessed September, 2014.

(3) Brookman, B., (1998) Guidelines for the In-House Production of Reference Materials - Version 2. National Measurement System - Chemical and Biological Metrology. Accessed September, 2014.

(4) ISO GUIDE 30 (1992) Terms and definitions used in conjunction with reference materials. International Organization for Standardization, Geneva, Switzerland. Accessed September, 2014.

(5) Roelandts, I., and Gladney, E. S. (1998) Consensus values for NIST biological and environmental standard reference materials. *Fresenius' J. Anal. Chem.* **360**, 327-338.

(6) Paule, R. C., and Mandel, J. (1982) Consensus values and weighting factors. *J. Res. of National Bureau of Standards* **87**, 377-384.

(7) Schiller, S. B., and Eberhardt, K. R. (1991) Combining data from independent chemical analysis methods. *Spectrochim. Acta* **46B**, 1607-1613.

(8) DerSimonian and Laird (1986) Meta-analysis in Clinincal Trials, *Controlled Clinical Trials*, **7**, 177-188

(9) Heckert, A. (2006)Consensus mean. Available from National Institute of Standards and Technology website. http://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/consmean.htm

(10) Keith, L. H., Crummett, W., Deegan, J., Libby, R. A., Taylor, J. K., and Wentler, G. (1983) Principles of Environmental Analysis. *Anal. Chem.* **55**, 2210-2218.

(11) Bevington, P. R. and Robinson, D. K. (1983). Data Reduction and Error Analysis for the Physical Sciences, 3rd edition pp 195-197. McGraw-Hill, New York.

(12) ISO GUIDE 31(2000) Reference materials - Contents of certificates and labels, International Organization for Standardization. Accessed September, 2014.

(13) Thompson, M., and Wood, R. (1993). The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories. *Pure Appl. Chem.* **65**, 2123-2144.