Anal Bioanal Chem (2007) 389:219–229
DOI 10.1007/s00216-007-1366-0

ORIGINAL PAPER

Reference materials to evaluate measurement systems for the nutrient composition of foods: results from USDA's National Food and Nutrient Analysis Program (NFNAP)

Katherine M. Phillips · Wayne R. Wolf · Kristine Y. Patterson · Katherine E. Sharpless · Joanne M. Holden

Received: 3 April 2007 / Revised: 7 May 2007 / Accepted: 11 May 2007 / Published online: 22 June 2007

Abstract Over a 6.5-year period a total of 2554 values were reported by nine laboratories for 259 certified or reference nutrient concentrations in 26 certified reference materials (CRM) submitted to contract laboratories, blinded, as part of the qualifying process for analytical contracts and in the routine sample stream as part of the National Food and Nutrient Analysis Program. Each value was converted to a Z'-score, reflecting the difference from the assigned value related to the combined expected analytical uncertainty plus the uncertainty in the CRM value. Z'-scores >|3.0| were considered unacceptable. For some nutrients (Na, folate, dietary fiber, pantothenic acid, thiamin, tocopherols, carotenoids, monounsaturated, and polyunsaturated fatty acids), >20% of Z'-scores were >|3.0|. For total fat, vitamin C, and niacin >25% of Z'-scores were >|2.0|. Components for which CRM data were best (more than 90% of Z'-scores <|2.0|) were Mg, P, Mn, Se, and vitamin B12. In some cases deviations from assigned values were not uniform across laboratories and materials. For Na almost all high Z'-scores were for low-Na matrices, suggesting analytical problems related to concentration.

Keywords Reference materials · Uncertainty · Accuracy · Food composition data

Disclaimer: Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, or the United States Department of Agriculture, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

K. M. Phillips (✉)
Biochemistry Department,
Virginia Polytechnic Institute and State University,
304 Engel Hall,
Blacksburg, VA 24061, USA
e-mail: kmpvpi@vt.edu

W. R. Wolf
Food Composition Laboratory,
USDA Beltsville Human Nutrition Research Center,
10300 Baltimore Ave., Bldg. 161, Rm 203C,
Beltsville, MD 20705, USA

K. Y. Patterson · J. M. Holden
Nutrient Data Laboratory,
USDA Beltsville Human Nutrition Research Center,
10300 Baltimore Ave., Bldg. 005, Rm 107,
Beltsville, MD 20705, USA

K. E. Sharpless
National Institute of Standards and Technology,
100 Bureau Dr., Stop 8390,
Gaithersburg, MD 20899, USA

Introduction

Certified reference materials (CRMs) play a critical role in validating the accuracy of nutrient data. Food matrix CRMs are intended to mimic "real" samples that an analyst might encounter and are used for several purposes: to facilitate testing of the accuracy of the entire assay system during development or implementation of an analytical method (e.g., from extraction through quantitation by instrumental analysis); for determination of whether a method is in control during routine use; for provision of traceability of values assigned to an in-house control material; or as a reference sample for assessment of inter-laboratory variability. A summary of currently available food CRMs with assigned values and uncertainty intervals for nutrient concentrations has recently been published [1].

The United States Department of Agriculture's continuing National Food and Nutrient Analysis Program (NFNAP) generates data for the United States Department of Agriculture (USDA) Nutrient Database for Standard Reference [2] based on nationwide statistical sampling and chemical analysis of nutrient composition of key foods [3]. CRMs have been used extensively in the NFNAP to monitor the accuracy of nutrient assays for a wide range of foods [4]. Over a 6.5-year period (1999–2006) a total of 2554 values were obtained for over 100 different components, from nine contract analytical laboratories for 259 certified or reference nutrient concentrations in 26 CRMs. These CRMs were analyzed either as part of the qualifying process for a project contract or from samples included for quality control in the routine sample stream [4].

The goal of this paper is to present the nutrient data for numerous currently available CRMs [1] that were analyzed during the NFNAP. This dataset differs from existing published analytical results for food-matrix CRMs that are exclusively from inter-laboratory efforts to generate assigned or consensus values for certificates (for example Refs. [5–10]), from studies performed to validate new analytical methods (for example Refs. [11, 12]), or for selected individual components in CRMs to validate analytical data in specific single-laboratory studies. In contrast, results from the NFNAP were obtained primarily from major commercial laboratories in the US, that operate on a fee-for-service basis, to produce nutrient data on a large scale analogous to the manner in which samples would be customarily analyzed at commercial laboratories, using the methods and quality control routinely implemented for clients (e.g., the food industry for food labeling). Furthermore, collection of these data spanned a much longer time frame (years as opposed to weeks), and the samples did not receive special attention to detail and calibration that would be expected in an organized characterization exercise. While the US Food and Drug Administration's Total Diet Study [13] has generated reports on selected elements (Ca, Cu, Fe, Mg, Mn, P, K, Se, Na, Zn) in some CRMs, most of those analyses were not performed at commercial laboratories, and are generally limited for organic nutrients.

The presentation of data from the NFNAP in this paper is intended as an initial evaluation of the overall quality of production-scale analytical measurement systems for nutrients in various matrices, to focus attention on areas needing improvement. This dataset was neither designed to be a balanced statistical study as are those performed prospectively by National Metrology Institute laboratories (such as the international studies by organizations such as the Consultative Committee for Amount of Substance – Metrology in Chemistry (CCQM; http://www.bipm.fr/en/committees/cc/ccqm/)) to evaluate the state of the art for specific analytical measurements, nor was it a proficiency test of the laboratories or a round robin or ring trial. Almost all proficiency test programs focus on reproducibility among the participants rather than accuracy per se. In contrast, the results presented in this paper represent extensive "real-life" data from CRM samples of well-characterized composition that can provide insight on accuracy and variability of routine analytical data from commercial laboratories.

Materials and methods

Procurement of CRMs and their use during the National Food and Nutrient Analysis Program have been described previously [4]. Briefly, laboratories bidding on NFNAP contracts initially analyzed CRM samples and the results were evaluated to qualify laboratories for award of the analytical contracts. At that time some facilities were disqualified for selected nutrients on the basis of previously determined criteria for acceptance. Laboratories awarded contracts represented those that generated satisfactory results for the test samples. They were all major commercial laboratories that routinely perform analyses for the food industry (although that was not a requirement for the contract). Subsequently, CRMs were submitted to these contract laboratories batched with test samples of similar matrix within the routine sample stream, and assayed in singlicate. The dataset used in this study includes values from both the qualification process for laboratories that did and did not qualify, and from the routine sample runs.

All of the laboratory data were converted when necessary to the same basis as the certificate values. In the case of dry CRMs with values assigned on a dry-mass basis, laboratory values were adjusted using the assayed moisture concentration determined by drying to a constant mass under vacuum [14]. Fatty acid concentrations were converted to the same units reported in the certificate of analysis for a given CRM (e.g., triglyceride, free fatty acid, or methyl ester) using the respective molecular masses of the methyl esters, acid, and triglyceride forms [15]. If *cis* and *trans* isomers of unsaturated fatty acids were itemized but the certificate value was for the undifferentiated form, the laboratory data were summed to give an undifferentiated value; when isomeric forms were reported, the laboratories were specifically queried to ensure that their designation of the fatty acid was correctly identified for comparison. In some cases the certificate gave isomer values but the laboratory gave undifferentiated data, and no comparison was possible. Data for the following major dietary fatty acids were evaluated: C10:0, C12:0, C14:0, C16:0, C16:1, C18:0, C18:1, C18:2, C18:3, C20:0, C20:5, C22:6, C22:0, C22:5, C24:0. Fiber in all cases was

determined by enzymatic/gravimetric analysis (e.g., Ref. [16]), and analytical results were compared only to certificate values determined by the same methodology. Cases of obvious reporting errors, which did in fact occur, were omitted. These included nitrogen concentrations reported as protein, and values that were high or low by an order(s) of magnitude, suggesting an error in concentration units, data entry, or calculation (unless the values were in fact verified as correctly reported). While these errant data were excluded from this analysis, it is important to note that such problems did occur.

Each assayed nutrient concentration was converted to a Z'-score following a modification of the approach of Jorhem [17], as follows:

$$Z'Score = \frac{(X - Xc)}{\left(\left(\frac{\frac{Xc}{100}\times 2^{(1-0.5\log(Xc*f))}}{\sqrt{n}}\right)^2 + \left(\frac{U}{2}\right)^2\right)^{0.5}} \quad (1)$$

where X is the individual laboratory result, Xc and U are, respectively, the assigned value and uncertainty from the certificate of analysis for the CRM, f is the fractional factor (e.g., 10^{-6} for μg g^{-1}, 10^{-5} for mg/100 g) [18], and n is the number of replicate laboratory analyses (n=1 in the present study, based on single analysis of CRM samples in any given analytical batch). This formula expresses the difference from the assigned value in terms of the uncertainty in the assigned value and a factor for the expected uncertainty in the laboratory result, as suggested in the approach of Jorhem (Ref. [17], "Procedure 1.3"). The expected uncertainty in the laboratory result as approximated by the "Horwitz Curve" [18, 19] is the first factor in the denominator of Eq. 1, and was calculated using the CRM's assigned value. A Z'-score may be either positive or negative, reflecting either a higher or lower result compared to the assigned value.

The formula for Z'-score accounts for both the expected uncertainty from the laboratory [18] and the uncertainty in the CRM's assigned value and is a modification of the approach of Jorhem [17]. This calculation of a Z'-score is similar to the International Union of Pure and Applied Chemistry (IUPAC) Z-score widely used in laboratory proficiency testing programs, in which $Z=(x-X)/\sigma$ (with x being the individual laboratory result, X the exercise mean value, and σ the exercise standard deviation), with a Z-score of 1.0 representing 1 exercise standard deviation [20]. The recent publications by Jorhem [17, 21] have addressed methods for comparing certified values to experimentally determined results using a CRM's standard deviation to calculate a Z-score. In the NFNAP data set, it was not possible to obtain the standard deviation of the certified value unless that information was provided on a material's certificate, and this information was not provided for all the CRMs used. Because the uncertainty in CRM values is often described as a 95% confidence interval, which can be approximated as 2SD, U/2 was used in Eq. 1. Although the Horwitz function [18, 19] used to estimate the expected between-laboratory uncertainty does not provide a rigorous estimate [22], the guidance it provides on the relative variability of chemical measurements as a function of analyte level in well-controlled studies was considered fit for the purpose of the present study and is a widely accepted concept, derived using data from thousands of collaborative analytical studies [18].

Statistical analyses

Statistical analysis of the Z'-scores was performed for selected nutrients to test the significance of apparent differences between particular laboratories or CRMs, using an analysis of variance via the Proc GLM procedure, with means compared by Tukey's test with a significance level of α=0.05 (SAS Release 8.2 (TS2M0), SAS Institute, Cary, NC, USA).

Results and discussion

Over a 6.5-year period (1999–2006) a total of 2554 values from nine laboratories for 259 certified or reference nutrient concentrations in 26 CRMs were obtained. Most of the results were for NIST materials (2203), with fewer for BCR (182), AACC (128), and LGC (91) products. The evaluation and acceptance of data for NFNAP food samples analyzed along with the CRMs was based on the results for the CRMs along with other considerations, as reported previously [4]. When CRM results were unsatisfactory, the sample values were typically rejected and the samples were reanalyzed.

A number of issues had to be addressed before the laboratory data could be compared to the assigned values. There were many cases in which the CRM assigned value was on a dry basis and laboratory results had to be adjusted using the moisture concentration. The moisture corrections ranged from nearly 13% for one of the flour CRMs (BCR 121 Wholemeal Flour) to a little less than 2% for lyophilized mixed vegetables (BCR 485). Also, it was not infrequent that the nutrient form or units of measure differed from those on the certificate, e.g., retinol versus retinyl palmitate; α-tocopherol versus α-tocopheryl acetate versus "vitamin E"; vitamin D in international units (IU) versus mass fraction; fatty acids as triglycerides, free fatty acids, or methyl esters. Awareness of these issues was critical in this study, and is important to note for reliable use of CRMs in general.

In this study, the absolute values of the Z'-scores were evaluated, unless specifically noted. Tables 1 and 2 summarize the Z'-scores for all results. Generally a Z'-score

Table 1 Summary of Z'-scores[a] for reported nutrient concentrations analyzed in certified reference materials (CRMs)

Class	Nutrient	Total CRMs	Total labs	Total values	Count of 0 to |1|	Count of |1| to |2|	Count of |2| to |3|	Count of >|3|	Percent > |2|	Percent >|3|
Proximates	Moisture	11	7	118	82	22	9	5	11.9	4.2
	Protein	9	5	106	60	24	12	10	20.8	9.4
	Ash	11	5	107	55	26	11	15	24.3	14.0
	Total Fat	11	6	129	52	39	15	23	29.5	17.8
Carbohydrates	Sucrose	4	3	28	18	9	1	0	3.6	0.0
	Fructose[b]	2	3	15	0	1	1	13	93.3	86.7
	Glucose[b]	1	3	8	2	1	3	2	62.5	25.0
	Lactose[c]	1	2	6	1	1	4	0	66.7	0.0
	Maltose[c]	1	3	9	5	3	1	0	11.1	0.0
	Dietary fiber (total)	5	4	31	4	7	8	12	64.5	38.7
Minerals	Ca	10	5	106	50	32	13	11	22.6	10.4
	K	10	5	106	64	29	6	7	12.3	6.6
	Mg	9	5	105	90	12	2	1	2.9	1.0
	Na	11	5	124	68	21	8	27	28.2	21.8
	P	9	5	105	65	32	6	2	7.6	1.9
Trace Elements	Cu	8	5	90	50	30	3	7	11.1	7.8
	Fe	8	5	91	49	28	3	11	15.4	12.1
	Mn	8	5	74	66	4	2	2	5.4	2.7
	Se	4	4	21	13	6	2	0	9.5	0.0
	Zn	8	5	98	79	9	7	3	10.2	3.1
Vitamins, water-soluble	B1 (thiamin)	9	5	73	27	20	2	24	35.6	32.9
	B2 (riboflavin)	8	5	68	41	11	10	6	23.5	8.8
	Niacin	6	5	73	26	22	13	12	34.2	16.4
	B6	6	3	49	17	22	6	4	20.4	8.2
	B12	4	3	43	33	7	3	0	7.0	0.0
	Folate	5	5	47	8	20	6	13	40.4	27.7
	Pantothenic acid	4	3	49	17	9	7	16	46.9	32.7
	C	3	4	36	21	5	3	7	27.8	19.4
Vitamins, fat-soluble	Retinol	1	3	24	18	5	1	0	4.2	0.0
	D	3	1	3	1	2	0	0	0.0	0.0
	Tocopherols	3	4	42	13	10	5	14	45.2	33.3
	Carotenoids	2	4	47	13	10	3	21	51.1	44.7
Other	MUFA	3	4	36	7	9	10	10	55.6	27.8
	PUFA	3	4	69	14	21	16	18	49.3	26.1
	SFA	4	4	107	57	32	10	8	16.8	7.5
	Amino acids[d]	1	3	225	83	93	32	17	21.8	7.6
	Choline	1	1	4	3	0	0	1	25.0	25.0
	Xanthines[e]	1	3	16	6	2	2	6	50.0	37.5
	Cholesterol	3	6	66	39	19	4	4	12.1	6.1

[a] See Eq. 1 in text
[b] All values are for LGC7103 Sweet Digestive Biscuit (LGC Promochem; Teddington, UK)
[c] All values are for LGC7107 Madeira Cake (LGC Promochem; Teddington, UK)
[d] All values are for NIST SRM 2387 Peanut Butter (National Institute of Standards and Technology, Gaithersburg, MD, USA)
[e] All values are for NIST SRM 2384 Baking Chocolate (National Institute of Standards and Technology, Gaithersburg, MD, USA)

less than or equal to 2.0 is considered satisfactory, between 2.0 and 3.0 questionable, and greater than 3.0 unsatisfactory [20]. While it is not possible to fully analyze results for each nutrient in this paper, some general observations can be made. Considering components for which there were greater than 30 analytical values, those with less than 10% of the overall Z'-scores in excess of |2.0| were magnesium, phosphorus, manganese, selenium, and vitamin B12 (Tables 1 and 2). For total dietary fiber, carotenoids, and monounsaturated fatty acids (MUFA), more than 50% of Z'-scores were outside ±2.0. These nutrients and also Na, folate, pantothenic acid, thiamin, tocopherols, and polyunsaturated fatty acids (PUFA) had >20% of Z'-scores in excess of |3.0|. Across all data results for the water-soluble vitamins showed considerable deviation from the CRM assigned values (with all but vitamin B12 having greater

Table 2 Classification of Z'-score[a] results by nutrient, for cases with greater than 30 values

Percent of total Z'-scores outside ±2[b]				
Less than 10%	10% to 15%	16% to 25%	26% to 49%	50% or greater
Mg	Moisture	Protein	Total fat	***Total dietary fiber***
P	K	Ash	***Na***	***Carotenoids***
Mn	Zn	Ca	Vitamin C	***MUFA***
Se	Cu	Vitamin B2 (riboflavin)	Niacin	
Vitamin B12	Fe	Vitamin B6	***Folate***	
	Cholesterol	SFA	***Pantothenic acid***	
		Amino acids	***Vitamin B1 (thiamin)***	
			Tocopherols	
			PUFA	

[a] See Eq. 1 in text
[b] Nutrients for which greater than 20% of Z'-scores were outside ±3.0 are shown in ***boldface italics***

than 15% of values outside |2.0|), and deviations for fat-soluble vitamins (tocopherols and carotenoids) were even greater (45% to 51% of values outside |2.0|). Considering also the typically higher uncertainty in assigned values for these components [1], the data reinforce conventional analytical wisdom on the difficulty of measuring these analytes in foods (for example, the published evaluations of folate [23] and carotenoids [24, 25]).

Figure 1 illustrates the Z'-scores by nutrient, laboratory, and CRM. Formal statistical analysis by laboratory and CRM was not possible for all nutrients, because the dataset is retrospective and unbalanced with respect to laboratories and CRMs (in other words, these data were accumulated for CRMs used to monitor analyses for the NFNAP). Nonetheless, some trends can be observed and it was possible to apply an analysis of variance in selected data subsets as a preliminary assessment to identify cases where deviations might be laboratory or CRM-specific rather than suggestive of overall methodological problems.

Proximates

Generally Z'-scores fell within |2.0|. Moisture and protein had only a few outlying values. For ash, Laboratory A had a large number of low Z'-scores (10 of 54 values), while Laboratory E had several high values. Overall there were an unexpectedly large number of Z'-scores exceeding |3.0| for fat, with four of six laboratories having >12.5% outside ±3.0. Results from Laboratory E for fat in NIST SRM 1546 (Meat Homogenate) differed significantly from other laboratories' data ($p<0.001$).

Carbohydrates

Data for sugars were limited, but it is interesting that all Z'-scores for sucrose, lactose, and maltose were <|3.0|, but most results for fructose exceeded |3.0|. This trend was consistent for all laboratories. Because only two CRMs are represented it is uncertain whether the deviations suggest a methodological or CRM/matrix problem.

Results for total dietary fiber were scattered, with no overall difference among laboratories, but with relative standard deviations from 22% to 95% within laboratory. All of the facilities reported using AOAC standard methods that were enzymatic/gravimetric [16]. Some samples were assayed for soluble and insoluble fiber and those concentrations were summed to yield total fiber. In such cases, if soluble fiber was less than the limit of quantitation (typically 1.0 g/100 g), total fiber was likely underestimated. Laboratory D had the smallest standard deviation and had also modified the methodology to be able to use larger samples, which may have contributed to fewer of their Z'-scores exceeding |3.0|.

Minerals and trace elements

Z'-scores for magnesium, phosphorus, and selenium were almost entirely <|3.0|, with most <|2.0|. Potassium, manganese, and zinc had only a few outlying results.

The large variability in Z'-scores for calcium was somewhat surprising. Across all CRMs the mean Z'-score for calcium from Laboratory A was significantly higher than that from Laboratories B and C ($\alpha=0.05$), and more than 10% of the Z'-scores from Laboratories A, C, and E fell outside |3.0|, while all of the values from Laboratory D were within |2.0|.

Difficulties in iron determination (contamination and incomplete solubilization) are known and may explain the variable results for this nutrient (Fig. 1b). All but one of the 11 high values were for the same CRM (NIST SRM 2383 Baby Food). Compared to other CRMs with eight or more values reported (NIST SRM1546 (Meat Homogenate) and NIST SRM 2387 (Peanut Butter)), the Z'-score for this material was significantly different ($P<0.0001$). Because

the magnitude of the assigned values for these three materials is similar (0.844 to 1.64 mg/100 g), as are the uncertainty intervals (4% to 9% of the assigned value), it does not appear that deviations are related to concentration.

The numerous high Z'-scores for Na (21.8% >3.0) were hypothesized to be concentration-dependent. In many fresh fruits and vegetables sampled for the NFNAP the assayed Na content did not agree well with existing data, and corresponding results for low-Na CRMs that were included as reference materials also showed deviations from assigned values, even from laboratories generally reporting accurate data for higher-Na foods (USDA Nutrient Data Laboratory, Beltsville, MD, USA; unpublished data, 2002–2003). Figure 2 shows Z'-scores for Na categorized by concentration. CRMs with a relatively low Na content (less than 30 mg/100 g) clearly accounted for the overall picture of high Z'-scores (Table 1 and Fig. 1b), with Z'-scores for CRMs with Na greater than 400 mg/100 g falling almost entirely within ±2.0. Possible explanations are contamination of samples during analysis (e.g., leaching of Na from glass if glass vessels are used during sample digestion), or instrument optimization for measurement of higher Na levels, leading to bias at low concentrations. It is unlikely that contamination during repackaging after certification was the cause, based on experiments conducted on the moisture and Na content of subsamples of NIST SRM 2385 Slurried Spinach, which support both homogeneity of Na in this material and results in agreement with the certificate of analysis (Patterson, K. and Phillips, K., unpublished data). High Na results were also observed in a previous study of NIST SRM 2384, Baking Chocolate [8], which has an assigned Na concentration of 3.8 mg/100 g to 4.2 mg/100 g. NIST SRM 2384 was not analyzed for Na during the National Food and Nutrient Analysis Program; therefore no values represented in the present study correspond to this material. Care should be taken to select control materials with analyte concentrations that match those of the test sample, even in products that have a similar overall matrix. For the NFNAP, low-Na samples (e.g., fresh fruits and vegetables) were subsequently analyzed after lyophilization to increase the Na content of the analytical aliquot.

Water-soluble vitamins

Although vitamin B12, is considered difficult to measure because of low (microgram) levels, results in this data set represented one of the best cases (<10% of Z'-scores greater than |2.0|). Riboflavin is regarded as more problematic than other water-soluble vitamins because of solubility and stability issues, but in the present evaluation data were generally acceptable, except for several values >3.0 for NIST SRM 2385 (Slurried Spinach) and NIST SRM1546 (Meat Homogenate).

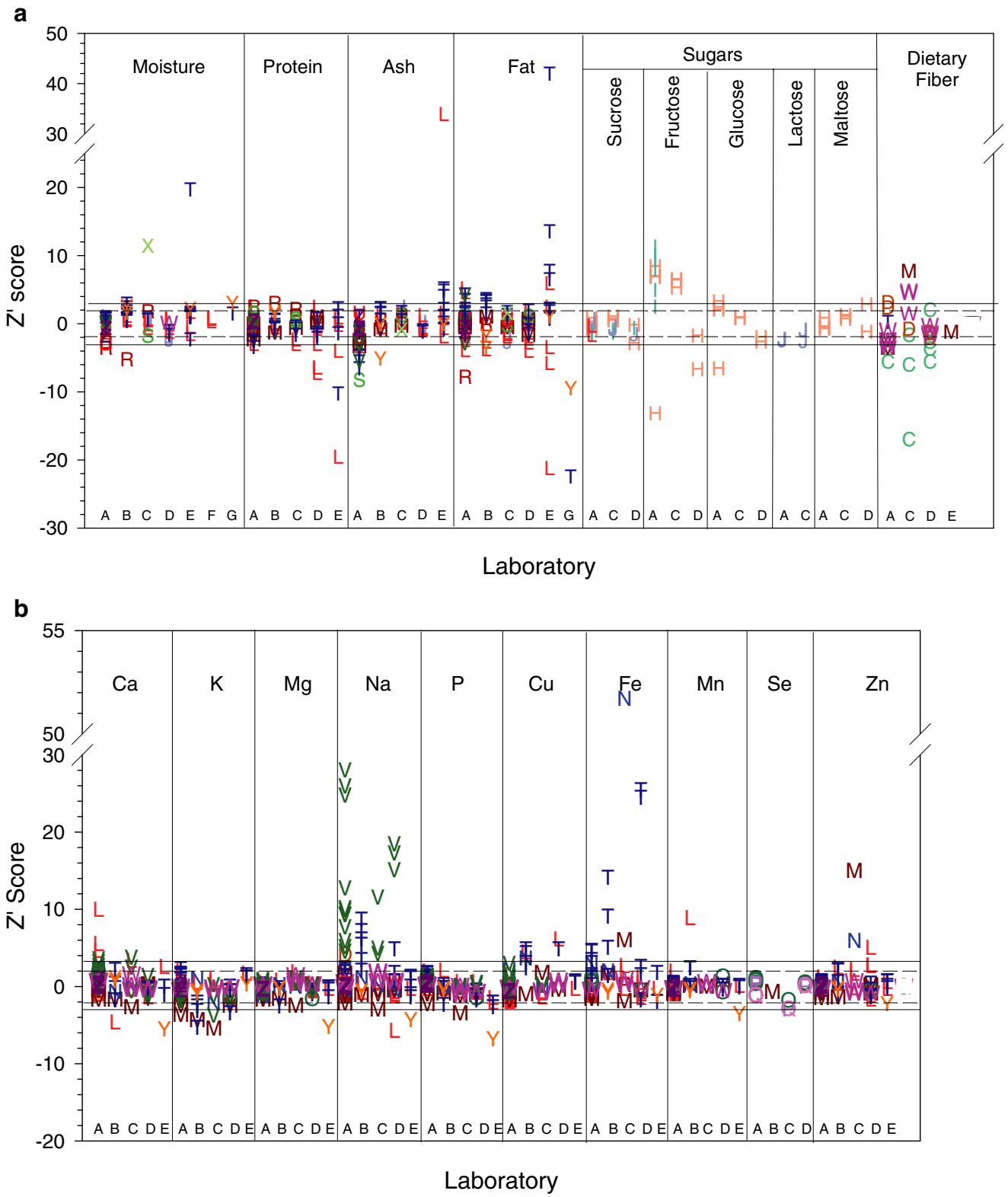
Existing literature suggests that methods for thiamin and niacin are generally satisfactory, but in the present study high Z'-scores indicated difficulties. For thiamin, many Z'-scores were within ±3.0, but nearly all results were <−3.0 for one material (NIST SRM 1546 Meat Homogenate) and differed significantly from those for other CRMs ($P<0.0001$) across all laboratories. These low scores could reflect an overall low bias with routine measurement of thiamin in meat or an issue with the assigned value for the CRM. Since these values for thiamin generated by the contract analytical laboratories over the six-year period were consistent, a reassessment of thiamin measurements in this matrix could be worthwhile (NIST plans to investigate this apparent discrepancy).

Niacin showed some high and low Z'-scores, with most laboratories reporting a few values >|3.0|. These data show that although niacin is generally a more straightforward determination, in practice some significant outliers can occur.

Vitamin B6 is not considered particularly difficult to analyze, although it does occur as three vitamers so it is important that an analytical method adequately measure all forms. Overall Z'-scores for vitamin B6 were good; however, 25% (7 of 28 values) from Laboratory A exceeded |3.0|, with those for BCR485 (Freeze-Dried Mixed Vegetables) and NIST SRM 1546 (Meat Homogenate) tending to be high, and those for NIST SRM 2383 (Baby Food) tending to be low. NIST SRM 2383 is a mixed food matrix, and given the assigned values of 0.151, 0.13, and 0.48 mg/100 g for NIST SRM 2383, NIST SRM 1546, and BCR 485, respectively, the high Z'-scores do not seem to be related to matrix or concentration, though there are not enough data to draw firm conclusions.

For pantothenic acid, Z'-scores were scattered, with relatively many exceeding |3.0|. Most outlying values were

Fig. 1 Z'-scores[a] for reported nutrient concentrations in certified reference materials (CRMs) (Table 1), by laboratory and nutrient. **a**. Proximates and carbohydrates. **b**. Minerals and trace elements. **c**. Vitamins. **d**. Fatty acids, amino acids, and other components. *Solid lines* indicate ±3.0 Z' and *dashed lines* indicate ±2.0 Z'. (See Table 2 for total number of values for each nutrient).. **A** BCR121 (Wholemeal Flour); **B** BCR122 (Margarine); **C** BCR382 (Wheat Flour); **D** BCR383 (Freeze-Dried Green Beans); **E** BCR421 (Milk Powder); **F** BCR431 (Freeze-Dried Brussels Sprouts); **G** BCR485 (Freeze-Dried Mixed Vegetables); **H** LGC7017 (Sugar Confectionery); **I** LGC7103 (Sweet Digestive Biscuit); **J** LGC7107 (Madeira Cake); **K** NIST1544 (Total Diet); **L** NIST1546 (Meat Homogenate); **M** NIST1548a (Freeze-Dried Mixed Diet); **N** NIST1549 (Non-Fat Milk Powder); **O** NIST1567a (Wheat Flour); **P** NIST1568a (Rice Flour); **Q** NIST1577b (Freeze-Dried Bovine Liver); **R** NIST1846 (Infant Formula); **S** NIST1946 (Fish Tissue); **T** NIST2383 (Baby Food): **U** NIST2384 (Baking Chocolate); **V** NIST2385 (Slurried Spinach); **W** NIST2387 (Peanut Butter); **X** NIST8435 (Whole Milk Powder); **Y** VMA195 (Cereal); **Z** VMA399 (Cereal). See Ref. [1] for description of CRM suppliers, assigned values and uncertainty intervals for each material



from Laboratory A for NIST SRM 1546. Interestingly, for pantothenate "the microbiological methods are validated for the free forms of pantothenic acid only... for total pantothenic acid an enzymatic digestion ... is recommended... " [26]. The precise methodology used by commercial laboratories is not specified, but it is possible that variations in the conditions of microbiological assays used for these determinations differed, similar to the case with folate [23].

Although analytical issues with the stability and form of vitamin C are well documented [27–29], *Z'*-scores for vitamin

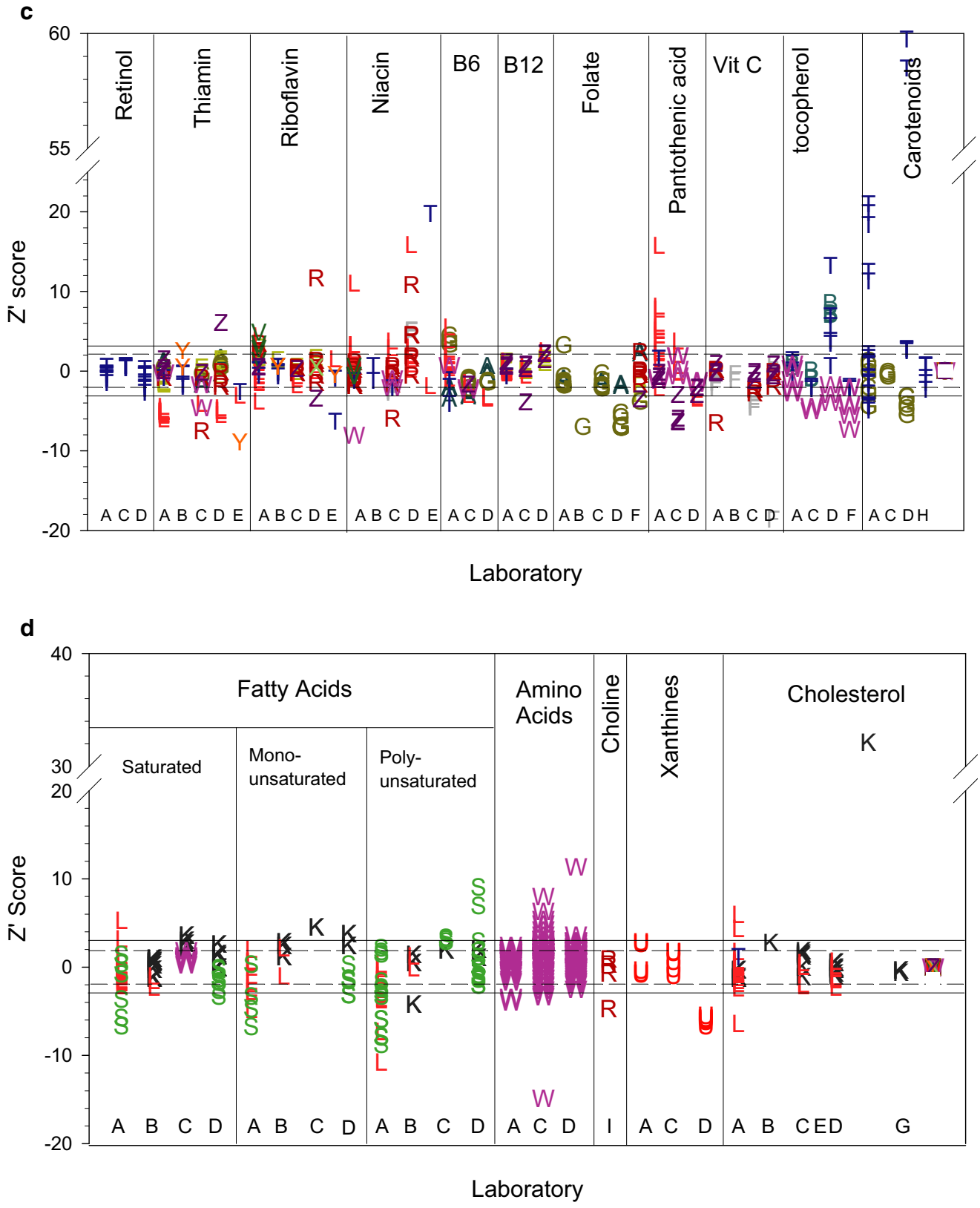


Fig. 1 (Continued)

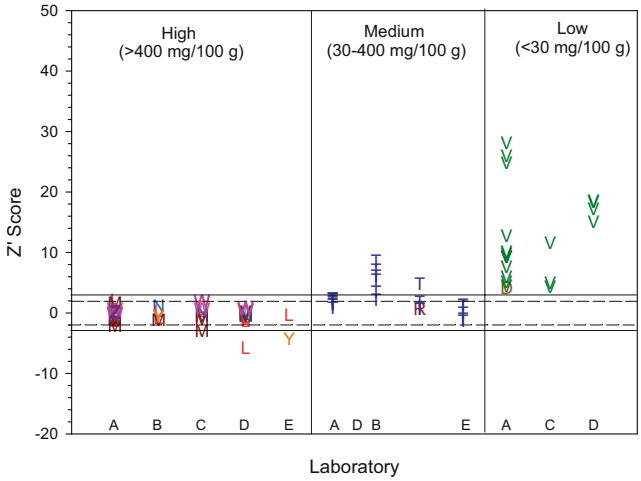


Fig. 2 Na Z'-scores[a] by concentration range. *Solid lines* indicate ±3.0 Z' and *dashed lines* indicate ±2.0 Z'. **D** BCR383 (Freeze-Dried Green Beans); V NIST2385 (Slurried Spinach); **T** NIST2383 (Baby Food); **L** NIST1546 (Meat Homogenate); **R** NIST1846 (Infant Formula); **M** NIST1548a (Freeze-Dried Mixed Diet); W NIST2387 (Peanut Butter); **N** NIST1549 (Non-Fat Milk Powder); **Y** VMA195 (Cereal); **Z** VMA399 (Cereal); **I** LGC7103 (Sweet Digestive Biscuit). See Ref. [1] for description of CRM suppliers, assigned values and uncertainty intervals for each material

C were nearly all within |3.0|. Most of the results were for BCR431 (Freeze-Dried Brussels Sprouts) and VMA399 (Cereal). Except for one outlying value each from Laboratories A and D, all values >|3.0| were from Laboratory C for BCR431, indicating a problem with analysis of vitamin C in this matrix at that facility; the negative Z'-scores (<−3.0) suggest possible problems with extraction or stability.

Z'-scores for folate varied widely, with 27.7% exceeding |3.0|. It is important to also note that the uncertainty intervals for folate CRMs are relatively large (9% to 22% of the assigned values) [1], so that these deviations represent significant imprecision and inaccuracy in measurement of this nutrient. Difficulties with folate analysis, especially of non-fortified foods, have been discussed in detail elsewhere [23, 30]. Consistent with these previous reports, the present dataset reinforces that greater difficulty generally exists in the analysis of endogenous folate (e.g., in vegetables, BCR485) compared to fortified matrices (e.g., NIST SRM 1846 Infant Formula). However, it was interesting that the most precise Z'-scores (all less than |2.0|) were for data from Laboratory C for BCR485 (Freeze-Dried Mixed Vegetables) which contains only endogenous folate.

Fat-soluble vitamins

All retinol data were for NIST SRM 2383 (Baby Food) and Z'-scores fell mostly within ±2.0 from all three laboratories represented, indicating good inter-laboratory precision for this matrix.

Results for carotenoids were for two CRMS, BCR485 (Freeze-Dried Mixed Vegetables) and NIST SRM 2383 (Baby Food). For each material there were significant differences in the overall mean Z'-score by laboratory ($P<0.006$). Z'-scores from two facilities (Laboratories C and H, Fig. 1c) showed excellent agreement with assigned values, while those from others were highly variable with many out of range points, especially from Laboratory A. Accurate and precise quantitation of carotenoids is a complex process that depends on careful attention to validation of the assay for specific matrices and to the execution and maintenance conditions of analysis [24, 25]. Differences among foods in the extractability and stability of carotenoids and the presence of interfering components (e.g., that co-elute during chromatographic separation or absorb at the same wavelength in a spectrophotometric assay) could easily lead to incorrect results when a standard method is applied in a "one size fits all" production-mode approach to different foods.

Considering all results for α-tocopherol, Laboratory D had significantly higher Z'-scores ($P<0.005$). All three values for BCR122 (Margarine) from Laboratory D exceeded 3.0, and there were also many high Z'-scores from this facility for NIST SRM 2383 (Baby Food).

Fatty acids

Z'-scores for saturated fatty acids were generally centered around zero, except from Laboratory C. Values reported by Laboratory A spanned a wide range reported by Laboratory A. Variability in Z'-scores for PUFA maybe related to the lower concentration of some of these fatty acids in some of the CRMs (e.g., C18:3 and C20:5 in NIST SRM 1546).

Amino acids

The data (Fig. 1d) represent a total of 12 analyses of NIST SRM 2387 (Peanut Butter). Z'-scores were nearly all within |3.0| from two laboratories, but the third (Laboratory C) reported 14% of results outside ±3.0, with most being >3.0. Although not itemized by specific amino acid in Fig. 1d, further examination of the dataset revealed that Z'-scores for all but cystine, glutamic acid, methionine, proline, tryptophan, and valine were <|3.0|, and that most of the outlying Z'-scores were for methionine (six values) and proline (five values).

Cholesterol

Most of the results were for NIST SRM 1546 (Meat Homogenate) and NIST SRM 1544 (Total Diet) and except for a few values, Z'-scores were <|2.0|. The precise and accurate data from the six laboratories represented probably

reflect the increased attention to cholesterol measurements over the past several decades, including development and wide use of definitive analytical methodology and extensive monitoring programs for this important component, given its clinical significance [31, 32].

Other components

There were limited results and data for only one CRM each for choline (NIST SRM 1846 Infant Formula) and xanthines (caffeine and theobromine in NIST SRM 2384 Baking Chocolate). Z'-scores from Laboratory D were well below −3.0 for xanthines but well within range from the other two facilities. Choline was only assayed at one laboratory, and one of four values had a Z'-score less than −3.

Conclusions

An overview of a large number of analytical measurements for nutrients in food reference materials obtained during routine determinations by major US commercial laboratories has been presented. The results are intended to be a first evaluation to open a dialogue on this topic and to raise areas of analysis that appear to be important for further work. The overall dataset suggests that each of the laboratories performs some nutrient analyses especially proficiently, but that none of the laboratories is proficient in all analyses. Additionally, even when a laboratory generally returned accurate data, there were often several outlying values, indicating that random errors do occur. Such deviations may not be revealed if control samples are not submitted with test samples as in the NFNAP [4]. Therefore, care should be taken by clients of nutrient analysis laboratories to validate datasets using external measures such as including blinded CRMs or other matrix-matched control materials.

While it was not possible due to the unbalanced dataset to perform rigorous statistical analyses, including possible effects of CRM and laboratory, results for a group of nutrients (fiber, fructose, Na (at low concentrations), niacin, thiamin in meat (NIST SRM 1546), folate, pantothenic acid, PUFA) suggest that a further look at these measurement systems might be warranted, because none of the laboratories were able to generate data with Z'-scores consistently $<|3.0|$. In contrast, nutrients for which most results fell within assigned limits might be those that have been the subject of intense clinical and nutritional interest and therefore involve methodology that has received significant attention from the analytical community, such as cholesterol. Higher quality data for some food components, such as proximates and some minerals, may reflect long-standing and rugged methods for frequently analyzed components in a wide range of matrices. Subsequent research will focus on analysis of nutrient data (>25,000 values) for a large suite of food-control materials developed for and used during the NFNAP [4], which will enable more specific evaluation of precision among laboratories and food matrices, and an assessment of how precise nutrient data must be for particular applications.

Acknowledgments The statistics advice and assistance on statistics by David Duewer, of the NIST in developing Eq. 1 for Z'-scores used in this study is acknowledged. This study was conducted as part of specific cooperative agreement #Y1-HV-8116-11 between the United States Department of Agriculture (USDA) Nutrient Data Laboratory and Virginia Polytechnic Institute and State University, with support from the National Heart, Lung, and Blood Institute, National Cancer Institute, the National Institute for Dental and Craniofacial Research, the Office of Dietary Supplements, and numerous other Offices and Institutes through the interagency agreement #Y1-HV-8116 between the National Institutes of Health and the USDA.

References

1. Phillips KM, Wolf WR, Patterson KY, Sharpless, KE, Amanna KR, Holden JM (2007) Accred Qual Assur, in press, DOI 10.1007/s00769-007-0257-6
2. US Department of Agriculture, Agricultural Research Service. USDA National Nutrient Database for Standard Reference., US Nutrient Data Laboratory, Beltsville MD. http://www.ars.usda.gov/Services/docs.htm?docid=8964
3. Pehrsson PR, Haytowitz DB, Holden JM (2003) J Food Comp Anal 16:331–341
4. Phillips KM, Patterson KY, Rasor AS, Exler J, Haytowitz DB, Holden JM, Pehrsson PR (2006) Anal Bioanal Chem 384:1341–1355
5. Ollilainen V, Finglas PM, van den Berg H, de Froidmont-Görtz I (2001) J Agric Food Chem 49:315–321
6. Sharpless KE, Schiller SB, Margolis SA, Thomas JB, Iyengar V, Colbert JC, Gills TE, Wise SA, Tanner JT, Wolf WR (1997) J Assoc Off Anal Chem Int 80:611–621
7. Sharpless KE, Gill LM, Margolis SA, Wise SA, Elkins E (1999) J Assoc Off Anal Chem Int 82:276–287
8. Sharpless KE, Thomas JB, Nelson BC, Phinney CS, Sieber JR, Wood LJ, Yen JH, Howell DW (2002) J Agric Food Chem 50:7069–7075
9. Sharpless KE, Phinney CS, Wood LJ, Yen JH, Howell DW (2003) J Agric Food Chem 51:6745–6751
10. Welch MJ, Colbert JC, Gill LM, Phinney CS, Sharpless KE, Sniegoski LT, Wood LJ (2001) Fresenius J Anal Chem 370:42–47
11. Finglas PM, Wigertz K, Vahteristo L, Southon S (1999) Food Chem 64:245–255
12. European Committee for Standardization TC275, http://www.cenorm.be/CENORM/BusinessDomains/TechnicalCommitteesWorkshops/CENTechnicalCommittees/Standards.asp?param=6256&title=CEN%2FTC+275
13. Pennington JAT (2000) J Food Comp Anal 13:539–544
14. Association of Official Analytical Chemists (1990) Official methods of analysis of the Association of Official Analytical Chemists, method 934.01, 15th edn. Association of Official Analytical Chemists, Washington DC
15. Sheppard AJ (1992) Lipid manual: methodology suitable for fatty acid-cholesterol analysis. William C. Brown Publishers, Dubuque IA

16. Association of Official Analytical Chemists (1995) Official methods of analysis of the association of official analytical chemists, methods 991.43 and 985.29, 16th edn. Association of Official Analytical Chemists, Washington DC
17. Jorhem L (2004) Accred Qual Assur 9:305–310
18. McClure FD, Lee J-K (2003) Computation of HORRAT values. J AOAC Int 86:1056–1058
19. Thompson M, Lowthian PJ (1997) J AOAC Int 80:676–679
20. IUPAC (1993) The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories. Pure Appl Chem 65:2123–2144
21. Jorhem L, Engman J, Schröder T (2001) Fresenius J Anal Chem 370: 178–182
22. Linsinger TPJ, Josephs RD (2006) Trends Anal Chem 25:1125–1130
23. Koontz JL, Phillips KM, Wunderlich KM, Exler J, Holden JM, Gebhardt SE, Haytowitz DB (2005) J AOAC Int 88:805–815
24. Oliver J, Palou A (2000) J Chromatogr A 881:543–555
25. Kimura M, Rodriguez-Amaya DB (1999) Arch Latinoam Nutr 49: 58S–66S
26. Devries JW (1993) Water soluble vitamins, chap 9. In: Sullivan DM, Carpenter DE (eds) Methods of analysis for nutrition labeling. AOAC International, Gaithersburg, MD
27. Moser U, Bendich A (1991) Vitamin C. In: Machlin LJ (ed) Handbook of vitamins. Marcel Dekker, New York, pp 214–232
28. Montano A, Casado FJ, Rejano L, Sanchez AH, de Castro A (2006) J Agric Food Chem 54:2206–2210
29. Johnston CS, Bowling DL (2002) J Am Diet Assoc 102:525–529
30. Konings EJM (2006) J AOAC Int 89:1–5
31. Joint Committee for Traceability in Laboratory Medicine, http://www.bipm.org/utils/en/xls/jctlm_listI.xls
32. Ellerbe P, Meiselman S, Sniegoski LT, Welch MJ, White VE (1989) Anal Chem 61:1718–1723