Variability of Estimates of Nutrients in Foods

William Horwitz
Food and Drug Administration
Washington, District of Columbia

The variability of the reported concentrations of nutrients in foods arises from differences in (a) sampling (which includes inherent growing, processing, and distribution factors as well as the physical removal of representative portions); (b) methodological factors (use of different methods of analysis, particularly for method-specific analytes such as "fibers"); (c) operational performance by analysts and laboratories; and (d) interpretation of results (statistical analysis, removal of outliers, categorizing). A major factor for improving the reports of nutrient analysis would be incorporating quality control for these variable factors into all investigations of food composition and including in the final manuscript statements of the quality control specifications used and the extent to which they had been met.

Although scientists realize that there is a certain amount of variability in their measurements, very few have any idea as to its magnitude. Regulatory agencies in particular are sensitive to this variability because, before taking expensive legal action, they wish to be very sure that they are right. Consequently, these organizations will repeat their measurements and even take additional samples to verify their initial findings, all on a case-by-case basis. As a minimum, analytical results in the Food and Drug Administration (FDA) are checked by a second analyst before legal action is approved; in some cases a second laboratory is also used. These replicate measurements are rarely identical, and through long experience regulatory officials may assign "working tolerances" of how much allowance they will make for these differences. In many cases, the party responsible for the goods under investigation can be expected to report results which are at variance with those of the regulatory agency.

Initially, much of this variability was ascribed to methodology. Certainly different methods of analysis could be expected to give somewhat different results. To remove this potential source of variability, regulatory chemists over the past century evolved a verification system which required that any method used for enforcement must be validated by an interlaboratory study to demonstrate the performance characteristics of the method (1). Such an interlaboratory study is conducted by submitting a set of identical, homogenous test samples to a group of typical laboratories for analysis as unknowns. The final results must show acceptably low variability. "Acceptably low" in this context means that the results are usually close to each other and that this pattern is consistent with historical performance. Historical performance has usually been based upon the experience of individual laboratories and is usually summarized by a statement such as, "the results from proximate analysis should agree within plus or minus a few percent and the results on trace nutrients should agree within ±10%." With sufficient experience, laboratories set up control limits on a statistical basis so that, for example, 5% of the values generated by the measuring system can lie outside of the boundaries formed by plus and minus 2 standard deviations from the mean. When chemical results do not appear to follow a normal distribution or when considerably more than 5% of the data points transgress these boundaries, an investigation of the source of the excessive variability is warranted.

At the request of the Food and Agriculture Organization (FAO)/World Health Organization (WHO) Food Standards Program (Codex Alimentarius) Committee on Food Labeling, we in FDA have been reviewing the precision of the methods of analysis available for declaring on labels the nutrient content of food. We realized immediately that analytical measurements contribute only one part of the apparent overall variability of the nutrient content of food. The other

important components, and in many cases the major components, are (a) the sample, defined to include natural, intrinsic variation in the food, and (b) analytical performance, defined to include method, laboratory, and analyst operations, as shown in Figure 1.

Figure 1. The major sources of variability of nutrient data and extreme examples from each source.

Source of Variability		
SAMPLING	METHODS	PERFORMANCE
Arises from the heterogeneity of the commodity	Different methods respond to different components differently	Analysts perform differently from Training Experience Environment
COMMODITY AND LOT SPECIFIC	METHOD SPECIFIC	LABORATORY AND ANALYST SPECIFIC
Extreme Examples:		
(low): Soybean oil	Iron	Weight
(High): Frozen Dinner	Dieatary Fiber	Infrared spectroscopy

Usually confounded (inseparable)

If analytical error is less than 1/3 the sampling error, lowering analytical error is of little importance in reducing total erro -- Youden

Sampling

The term "sample" will be used here in a very broad sense as covering all variability arising from agricultural, processing, and distribution factors, as well as including the selection and removal of representative portions from representative lots. This broad usage must be carefully distinguished from the narrow attribute of mere removal of representative portions of a food from a specific lot. We will use terms such as a "sampling variability" when referring to the broad concept and "sample error" or "simple sampling" when referring to the narrow concept. The indiscriminate selection of the results from reports of analysis of samples purporting to broadly represent a food is undoubtedly one of the main contributors to variability in database records.

There is not much that can be done with sampling. To obtain a truly representative sample of any important commodity would require designing a sampling plan whose implementation would consume more resources than would be available for an entire database project. Undertaking a suitable sampling plan for a commodity that represents only a fraction of a percent of the food intake of a population is too unimportant to even consider. Consequently many compromises must be made at this point, the chief one of which is that you must take whatever data happen to be available. The best that can be done, in the absence of experimental data, is to estimate (guess), on the basis of experience, the variability that is likely to be encountered from the inherent differences arising from agricultural, processing, and distribution sources. Then the data acceptance process must be managed to ensure that the input data are maintained in statistical control as determined on the basis of the historical input. Data outside of the established control limits require an investigation as to the occurrence of possible mistakes or blunders, which are very difficult to discover after the fact. One helpful factor is that nature does not tolerate gross discrepancies; experience and common sense teach what values do not belong to a category. Therefore, removal of out-of-line data is not the difficult decision-making process that so often occurs in handling data when no information is available to assist in indicating the likely values.

However, even if sampling is confined to the narrow definition of removal of a representative portion from a lot, many papers purporting to supply nutrition data give little information regarding this type of "simple sampling"; therefore, it is impossible to provide a realistic estimate of the sampling variability either with time or with the characterizing, descriptive parameters of the food.

In connection with the preparation of databases, it should also be kept in mind that if a food was originally analyzed for a non-nutritional purpose, the analytical results may be biased with respect to nutritional purposes. For example, regulatory authorities may collect samples because they suspect contamination from chemicals or from filth. They will collect nonrepresentative ("focused") samples intended to contain the contaminant rather than seek a typical sample of the food.

When a sample reaches the laboratory, the first task of the chemist is to reduce that sample both in bulk and in fineness to a manageable size. As a result of this operation, a homogenous mass should be produced which is not expected to introduce any further sampling errors. Consequently, any variability exhibited by this analytical sample is assumed to be entirely the result of analytical operations.

Analytical Variability

The major contributors to analytical variability are the inherent biases of the method and the errors introduced by the analyst in the application of the method to a particular material.

Method performance is usually estimated *a priori* by organizations such as the Association of Official Analytical Chemists (AOAC), which compile manuals of approved methods. These societies perform studies by distributing homogenous materials to laboratories to be analyzed as unknowns by the method being tested to determine the fundamental variability among laboratories when the method is used by typical chemists. The variability found in these studies is expected to reflect variability exhibited in actual practice, although it is well known that the variability shown during a method-performance trial is usually less than the variability found in actual practice.

A century-old record exists in the *Journal of the Association of Official Analytical Chemists* of method-performance trials of approved procedures for the analysis of foods. During the first half-century, the data were merely tabulated and the reviewers were left to draw their own conclusions. During the past quarter-century, statistics were used to summarize and to analyze the data, largely as a result of the stimulating lectures of the late Dr. William J. Youden, then of the National Bureau of Standards (2). However, because of the presence of outliers in all measurements and the lack of a standard procedure for examining them, a uniform procedure for reporting interlaboratory analytical data did not exist. Fortunately, in 1987, the International Union of Pure and Applied Chemistry (IUPAC) produced a protocol, designated "IUPAC-1987," for the design, conduct, and interpretation of interlaboratory method-performance (collaborative) studies (3). This protocol, which includes a standard outlier removal procedure, has been accepted by the AOAC and numerous other methods-standardizing organizations in the food field such as the International Dairy Federation (IDF), International Association for Cereal Science and Technology, American Oil Chemists' Society, and International Commission for Uniform Methods for Sugar Analysis.

We first applied this harmonized IUPAC protocol to many of the method-performance studies conducted by the AOAC and the IDF on milk products (4). These methods, particularly for solids, fat, and protein, have had the benefit of over a century of fine-tuning. For the purpose of comparing variability (precision) across several orders of magnitude, we have to use the relative standard deviation, RSD, which is simply the ordinary standard deviation divided by the mean and then placed on a percentage scale by multiplying by 100. The ordinary standard deviation is expressed in the same units as the mean and consequently varies directly with the mean; RSD is dimensionless and is independent of the units, %, g/100 g, decimal fraction, mg/L, etc., and is often independent of concentration.

Figure 2 shows the distribution of the among-laboratories relative standard deviations (RSD_R) as a function of concentration for all of the 673 individual data sets in the milk products database. The concentration, expressed as a decimal fraction (where 1% = 0.01) is plotted on the x-axis on a logarithmic scale so that all of the data can be presented in a single figure. Also, it is given as a negative function to "open up the zero". If the conventional type of graph with 0 at the origin of the x-axis was used, many of the low concentration data sets would accumulate near the y-axis, and we would be unable to see their relationship to decreasing concentrations. Distinct clusters are seen at certain concentrations which correspond to the 3% fat and 3% protein of milk (-log 0.03 = 1.6), the 30% fat of cream (-log 0.3 = 0.5), the 40% moisture of cheese (-log 0.4 = 0.4), and the 0.1% phosphorus in milk (-log 0.001 = 3). It is readily seen that the y-dimension of the clusters (variability as relative standard deviation) increases with decreasing concentration (which is equivalent to an increasing [negative] logarithm).

The dotted lower line in Figure 2 is a grand summary of the RSD_R values for over 6000 interlaboratory data sets that we have examined for all types of analytes from aluminum to Zoalene, by methods which range from classical gravimetric analysis to modern mass spectrometry, at concentration levels from pure materials (100%; C = 1.0) to residues and contaminants at a fraction of a part per billion ($C = 10^{-9}$), in solids, liquids, and gases, and in matrices (commodities) which include air, blood, cosmetics, drugs, feeds, fertilizers, ores, pesticides, tissues, and water, as well as foods. Table 1 gives some RSD_R values at useful

Table 1 - Typical among-laboratories relative standard deviations (RSD_R) as a function of concentration expressed as a decimal fraction and conventionally.

Concentration		RSD_R
Fractional	Conventional	(%)
1.00	100%	2
0.01	1%	4
0.0001	0.01%	8
0.000001 (10^{-6})	1 ppm	16
10^{-8}	10 ppb	32
10^{-9}	1 ppb	45

concentration levels, taken from the summary curve. The upper curve of Figure 2 is twice the summary curve values and represents what we consider the upper limit of acceptable precision for interlaboratory studies. This upper curve is a ceiling, based upon our review of the RSD_R values of methods that have been accepted over the past century as approved methodology by the AOAC and other organizations.

All the points and curves on the figures and in the discussion should be considered as "fuzzy"; i.e., all averages, values, and parameters are surrounded by confidence intervals whose width depends upon the desired level of confidence of being right, or its converse, the acceptable level of risk of being wrong. An occasional value beyond the limits can be tolerated; having many values lying beyond the boundaries, however, calls for an investigation as to the cause.

An important aspect about this general curve is that it represents a first approximation, which is useful in the absence of overriding information. Figure 2 shows that RSD_R values of method-performance studies for milk products are somewhat better than those for run-of-the-mill data sets. When we look at the corresponding values from proficiency studies, in which the analyst is not restricted as to the method to be used, we often find that the RSD_R values for the data sets are somewhat worse than those for the method-performance studies. Another important point is that RSD_R refers to the among-laboratories precision. All chemists think that they can do better than the performance shown by the curve, and they can. Within-laboratory precision values, designated as RSD_r, are roughly one-half to two-thirds of the among-laboratories values. It is always the other laboratories that are inflating the RSD_R values!

Similar figures applying to more restricted groups have also been prepared from method performance studies of the major nutrients in food -- protein, carbohydrates, and fat; for the supplementary analytes needed to obtain carbohydrates by difference -- ash, moisture, and fibers (5); and for the major mineral elements in food -- calcium, magnesium, phosphorus, potassium, and sodium (6). These graphs are shown in subsequent figures.

Figure 3 shows the precision of a very well-behaved analyte, protein, as a function of C. Each of the 208 RSD_R values for protein from the food database (5) is represented by an upper-case "P," and each of the 201 RSD_R values from the milk products database (4) is represented by a lower-case "m." All the milk values and many of the food values are below the typical lower curve. All but one of the values are below the upper limit. The variability of protein analyses in the concentration range of 1 to 100% can be characterized by an RSD_R of 2%, with most values within the range of 1-3%. An occasional value near 4% in a series is acceptable within the 1-100% concentration range, but having many values above 4% is not acceptable. Within-laboratory variabilities as measured by RSD_r are about one-half of the RSD_R values. This range of 1-3% for RSD_R of protein analyses is further confirmed by the results from laboratory proficiency programs that have been conducted by the American Association of Cereal Chemists and by the American Oil Chemists's Society since the 1920s.

These results may be contrasted with the corresponding among-laboratories results for carbohydrates, fat, and fiber shown in Figure 4. The typical and limit lines are the same, of course, as in Figure 3, but in Figure 4 many individual values are above the upper limit line. Some of the RSD_R values approximate 100%. Many points have RSD_R values above 25% and many of these points are at concentration levels below about 10%. In Figure 5 we show just the fat results, omitting 23 points with RSD_R values above 25% at the concentration level below about 3% fat, "F," in nonmilk foods. Note that all the values from milk products, "f," in Figure 5 are below the "typical" line. The 23 omitted values cannot be considered as being in statistical control, whereas the values for fat in milk products are in excellent control. An important contributor to the high variability of the results for this analyte, fat, as well as those for moisture, ash, and fiber-related components, is the use of too small a test portion (colloquially but incorrectly called "sample weight"). If the test method is applied in such a manner that less than 50 mg of volatiles (moisture) or residue (ash, fat, and fiber components) must be weighed in the final measurement, high variability cannot be avoided. Some of the other conclusions of this review were that low-fiber foods containing less than about 5% fiber cannot be analyzed reliably, regardless of "improved" methodology; that because of disagreements on definitions (accuracy), it is impossible to obtain more reliable methods of analysis; that within-laboratory precision cannot predict among-laboratories precision; and that it is inappropriate to apply methods for fiber to products with very low "fiber" content (flour and rice at 0.5% and starches at 0.1%) and to fluid matrices such as milk and eggs.

Another group for analytes for which data are available are the major elements, shown in Figure 6 (6). Typically most values for these 5 major elements -- calcium, magnesium, phosphorus, potassium, and sodium -- are below the upper limit, but an appreciable

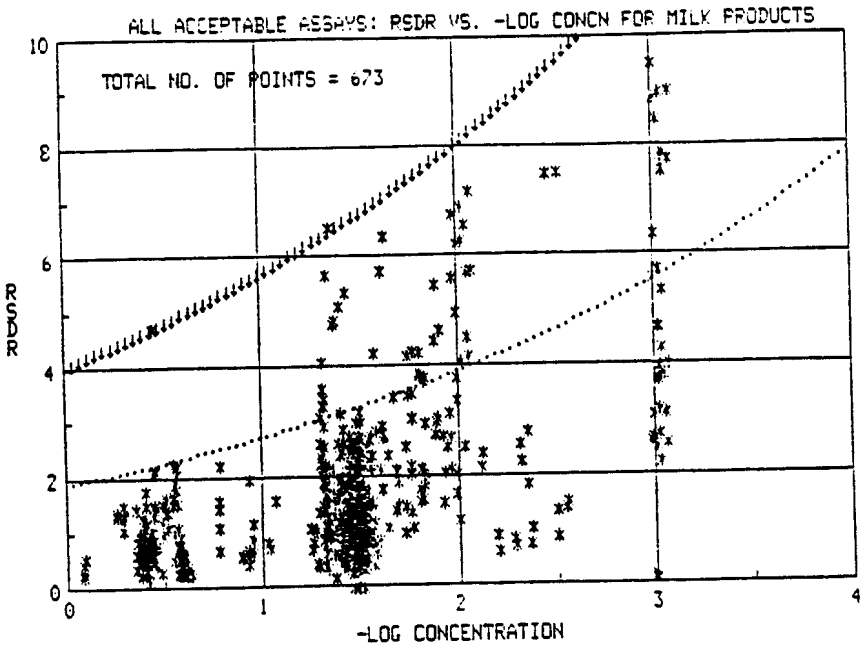

Figure 2. The among-laboratories relative standard deviations, RSD_R, for all 673 data sets of analytes (moisture/solids, ash, carbohydrates (by difference), fat, fiber-related, protein, individual sugars, and individual major elements) in milk products as a function of $-\log_{10}C$, where C is expressed as a decimal fraction. The lower line is represented by the equation, $RSD_R = 2^{(1 - 0.5 \log_{10}C)} = 2C^{-0.1505}$; the upper line is twice this curve and is considered the upper empirical acceptable limit for all analytes, independent of analyte, matrix, and method. (Figure 4 from ref. 4.)

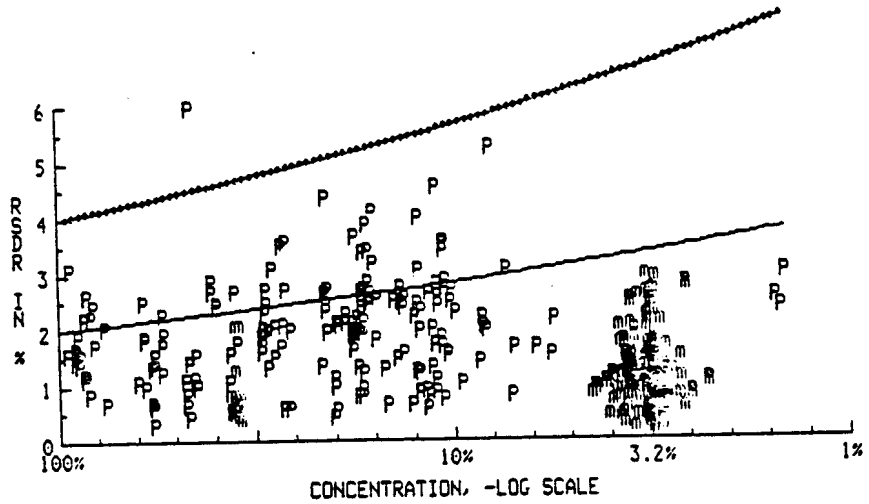

Figure 3. The among-laboratories relative standard deviations, RSD_R, for 208 data sets for protein in non-milk foods, P, and from 201 data sets for protein in milk products, m, plotted as a function of -log C as in Figure 2.

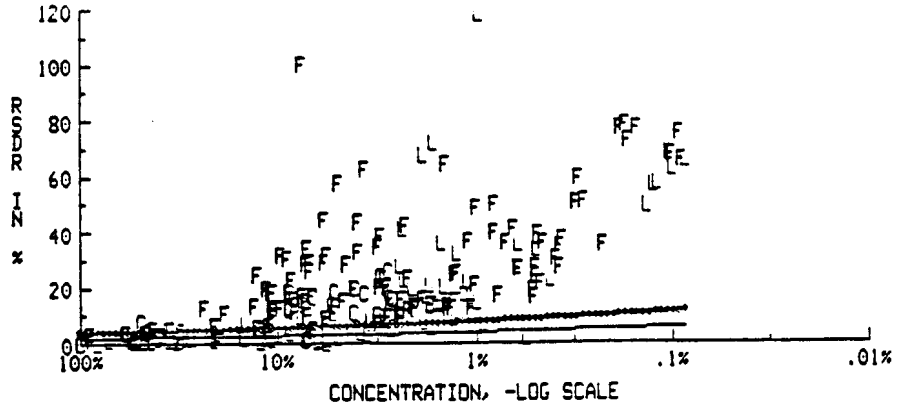

Figure 4. The among-laboratories relative standard deviations, RSD_R, for 107 data sets for fiber-related analytes, F; 60, carbohydrates, C; and 112, fat (lipids), L; all from foods, plotted as a function of -log C as in Figure 2.

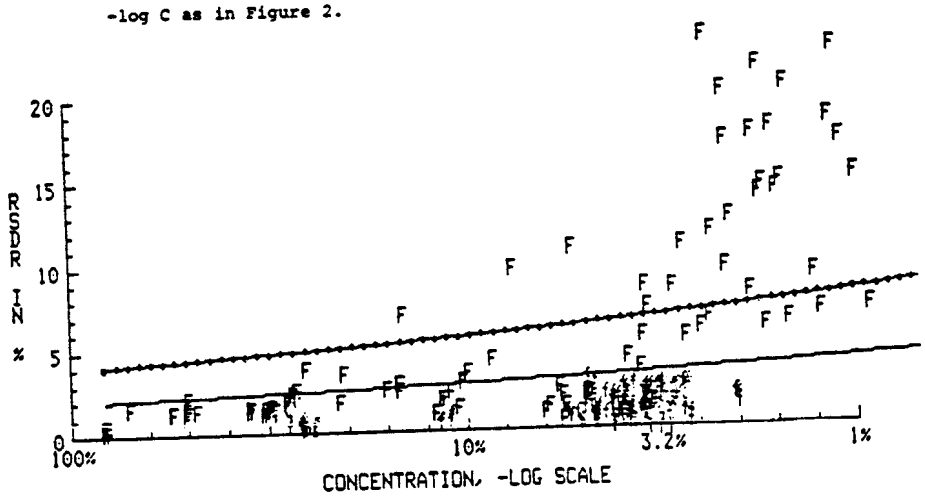

Figure 5. The among-laboratories relative standard deviations, RSD_R, for 89 data sets for fat in non-milk foods, F, and from 214 data sets for fat in milk products, f, plotted as a function of -log C as in Figure 2.

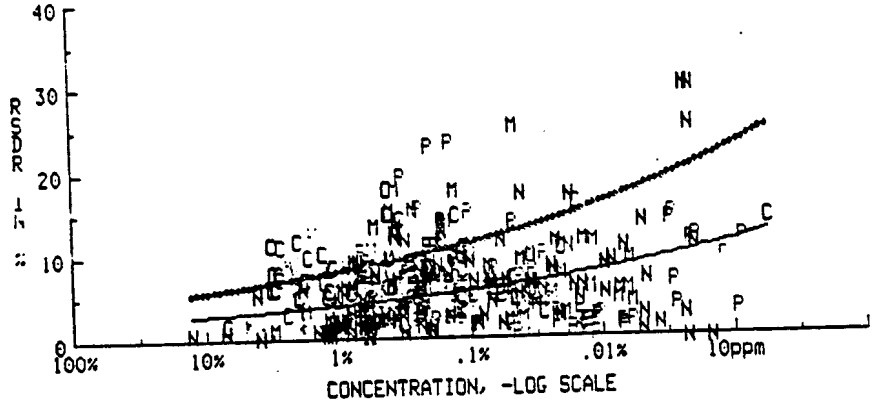

Figure 6. The among-laboratories relative standard deviations, RSD_R, for 89 data sets for calcium in foods, C; 48 magnesium, M; 125 phosphorus, P; 91 potassium, K; and 129 sodium, N; plotted as a function of -log C as in Figure 2.

fraction of values exceed the upper boundary of acceptability. This excessive variability is not confined to any particular element or to high or low levels. We interpret these data to indicate that food chemistshave not paid much attention to the necessity for quality control of their work but have depended upon familiarity with their procedures to produce the "correct" results. Furthermore, only now are certified reference materials becoming available that will permit food chemists to calibrate their performance against "true" values. With regulatory agencies now expecting the application of "good laboratory practices," with more consideration being paid to laboratory and analyst performance, and with laboratory accreditation being instituted in many countries, particularly within the European common market, much greater attention must be paid to the question of reliability of analytical results.

We are now extending our review of the available interlaboratory studies of methods for the minor and trace elements such as copper, manganese, selenium, and zinc; to vitamins; and to other compounds of nutritional interest such as cholesterol and amino acids.

Analyst Performance

Analyst performance is usually difficult to isolate because it is automatically tied up with methodology. In those cases where an attempt is made to separate performance from methods by having the same analyst use different methods, almost invariably the analyst shows much lower variability with the method that is in routine use in the laboratory. Most interlaboratory proficiency studies permit the analyst to use any method. Some studies are so designed that the method effect may be isolated. In most of these cases, when the study shows good control, the variability shown can be represented by the curves previously discussed. Rarely is performance any better than shown in the historical curves; frequently performance is considerably worse. The American Association of Cereal Chemists circulates proficiency materials to be analyzed for the enrichment ingredients niacin and thiamin, for which performance is considerably better than would be expected from the general precision curve. This is easily explained by the fact that well-standardized methods are used routinely by trained technicians who have all the operations on the homogenous test samples from uniform, standardized commodities under excellent control. For those laboratories that conduct these analyses sporadically, however, it can be expected that theirperformance would approach the general curves.

Experience with examination of analytical results indicates that out-of-control operation usually results from two main sources: (a) clerical errors in recording and transcribing numbers and in performing calculations and (b) improperly prepared standard solutions used for the calibration curves. Clerical and mathematical blunders can be controlled to a large extent by automating these operations. Incorrectly prepared reagents can be discovered through routine quality control procedures such as comparing a current calibration curve with historical data, and through the use of certified reference materials or even the routine use of "house" historical standards.

Routinely participating in professional proficiency studies and taking corrective action when problems are discovered is one means for maintaining the optimum performance of any laboratory. The best way is by randomly using blind in-house check standards and displaying the results as control charts like those often seen now in hospital laboratories. Such quality control operations should be a part of the normal operating budget and management control of laboratory functions.

Interpretation of Results

Compilers of databases are familiar with the pitfalls that can come from having to guess at the meaning of various aspects of the results of chemical analysis that authors have neglected to mention. These aspects may be the simple omission of factors used in the conversion of nitrogen to protein or a fundamental failure to indicate the steps that were taken to ensure the validity of the results. The need for validation of results is particularly important in food analysis because so many of the methods of analysis are method-specific (empirical). As indicated previously, in many cases the differences may not be very great; but in the case of dietary fiber such differences are the cause of endless polemics, even though the absolute differences are not very significant from the point of view of nutritional labeling. An absolute difference of 1% in dietary fiber in a product that contains 10% relative difference, will change the carbohydrates (by difference) by only 1% or 4calories/100 g. This difference is of the same order of magnitude as the RSD_R (10%) shown by the interlaboratory studies of methods for fiber-related analytes at the 10% concentration level, so reasonable allowances must be made for this variability.

The term "reasonable allowances" deserves further discussion. Regulatory officials are well aware of the variability produced by inherent processing factors and by chemical analysis. They make due allowance

for the inherent random variability from these sources. But the type of variability they make allowances for is the 2-sided variability that appears both above and below the target (labeled) amount. A series of analytical results that appear consistently on the low side of a label declaration typically is not a result of random variability but is rather (a) a symptom of a real deficiency of analyte, or (b) a bias in the performance of the analysis, which bias can be isolated by the use of the regulatory method (1) or by use of a certified reference material, if available. The fact that the value is within the confidence interval of the analytical and sampling error is not a defense against a deviation in the declaration of a nutrient. There is a clear danger inherent in the use of nutrient data obtained by the uncritical pooling of numbers found in the literature through the potential introduction of unsuspected biases. The resulting average or range may have little relationship to reality. The literature of food analysis is notorious for this type of error.

Discussion

It is difficult to generalize with respect to the relative importance of the factors contributing to the variability of results that may enter a database. Unless individual portions from a lot (technically known as increments) have been analyzed in the original work, and their deviations carried over to a database, the error ascribable even to "simple" sampling cannot be estimated. Very few investigations have been performed with the objective of determining this type of sample error, with all other variables -- method, operator, storage, etc. -- held constant. Most such investigations have been interested in other phenomena such as variety, geography, agricultural conditions, and production and distribution variables. In many investigations of nutrient content as a function of food production, the extent of sampling and analytical errors has not even beenconsidered.

Food chemists have never thought it important to assess the variability inherent in sampling for various reasons. Most commodities are purchased on an "as is" basis or by a contract. If the value is based on a specification, the laboratories and the methods of analysis are often specified. An arbitration procedure for settling economic disputes is usually a built-in requirement. In the case of promulgation of food standards by the FDA during the 1940s and 1950s, representative nationwide samples of many foods were obtained. The sampling and analytical variabilities were incorporated into the final standard by specifying a limit close to the minimum (or maximum, as the case required) found in commercial channels.

Even knowledge of "simple sampling" variability is important because often the method or the analyst is blamed for poor performance when actually the initial laboratory sample or the prepared analytical sample is responsible for the obvious variability. Unrepresentative sampling can introduce a large uncontrollable error, for which no allowance can be made *ex post facto.*

Methodology is probably not too important a contributor to the variability seen in nutrient databases except in a few very well-known cases such as dietary fiber. Even when different methods such as Kjeldahl nitrogen or Dumas nitrogen are applied to the same food, the difference in the final result is only a matter of a few tenths of a percent. The same is probably true when different times and temperatures are used in the determination of total solids, moisture, and ash. A small but significant difference may be apparent if chromatography is used for the separation and determination of individual sugars and the results are compared to the previously reported results for the same but combined group of "total reducing sugars, before and after inversion." The recent collaborative studies conducted by Tanner and Barnett (7) verify the applicability of the current AOAC methods for nutrients in foods in general to milk-based infant formulae. Practically all of the results from these studies bracket the general curve shown previously. The final phase of these studies, which are now under review for publication, includes apparently satisfactory method performance data for nutrients in foods for which there has been no official methods, such asiodine and vitamin K. One interesting and important aspect of the use of the general precision curve is the finding that the method for vitamin D are already quite acceptable. The precision data (among-laboratories), RSD_R = 20-40% in the literature do not reflect poor laboratory performance or poor methods but merely that this vitamin is being determined at levels of less than 1 ppb, for which an RSD_R of 45% is acceptable on historical grounds.

Probably the most important single factor responsible for variability in food constituent databases is the absence of quality control of the analytical work. Even if quality control existed in the reporting laboratory, statements to this effect are frequently missing in the final manuscripts. Only within the last few years have certified reference materials become available to practicing food chemists to provide an absolute reference point for their analyses (8) for the control of accuracy of those analytes that are not method-dependent.

One potential improvement in the situation, by making food chemists aware of the importance of quality control, is the application of the "expert sy

tem" concept developed by the U.S. Department of Agriculture (USDA), Beltsville Human Nutrition Research Center, to published nutrient data (9). This system evaluates the published reports and data and assigns a quantitative rating scale based on the number of laboratory samples reported, the validity of the analytical methods, the handling and documentation of the laboratory and analytical samples, the sampling plan, and the extent of analytical quality control, all important aspects of good analytical practices (10). The summation of the quality factors for each item results in an overall 3-factor confidence code (A (best), B, and C), indicating the relative degree of confidence the user can have in a grand mean value for the analyte in a food. The system has been applied to selenium (11) and copper (12) in food. In the case of copper, only 14% of the confidence codes for 218 foods for which reports had been examined rate A, 24% rated B, and 62% rated C (limited confidence due to limited data quantity and/or quality). In general, the large number of C ratings was an indication of the paucity of data. Although it is possible that quality control may have been performed as part ofthe published investigations but not mentioned in the manuscripts, it is more likely that little attention was given to this critical requirement. In the case of some important foods, the widely disparate literature values required implementation of confirmatory analyses by the USDA laboratory. The scheme now requires modification to take into consideration the current availability of certified reference materials as an alternative to the original use of multiple laboratories to evaluate analytical bias.

These findings suggest that not much confidence can be placed in data that reside in the literature unaccompanied by documentation that they were produced under controlled conditions. This situation is not confined to inorganic elements, for which certified reference materials of reasonable similarity to the foods of interest are available as a check on the correctness of the results. Holden and Davis (13) describe their experience in selecting an analytical contractor for the analytical phase of a nationwide survey to update the USDA nutrient data for eggs. The results from blind analysis of National Institute of Standards and Technology (NIST) Standard Reference Material (SRM) 1563-2, Cholesterol and Fat-Soluble Vitamins in Coconut Oil, were used in part to select a technically competent contractor. The certified values was 624 µg/g with an acceptable range of 601-674 µg/g; only 2 of the 5 commercial and university laboratories submitted results within the limits of acceptability set by NIST. The reported results of analyses arranged in increasing order were 218, 287, 607, 643, and 866. Replication of results within a laboratory is not an acceptable quality control technique. The triplicate determinations in all laboratories indicated good precision. Only the availability of a reference value salvaged this phase of the study. Even the mean of these results, 524 µg/g, was outside the acceptable range, and the relative standard deviation of the 5 values was an unacceptable 50%. Government con-

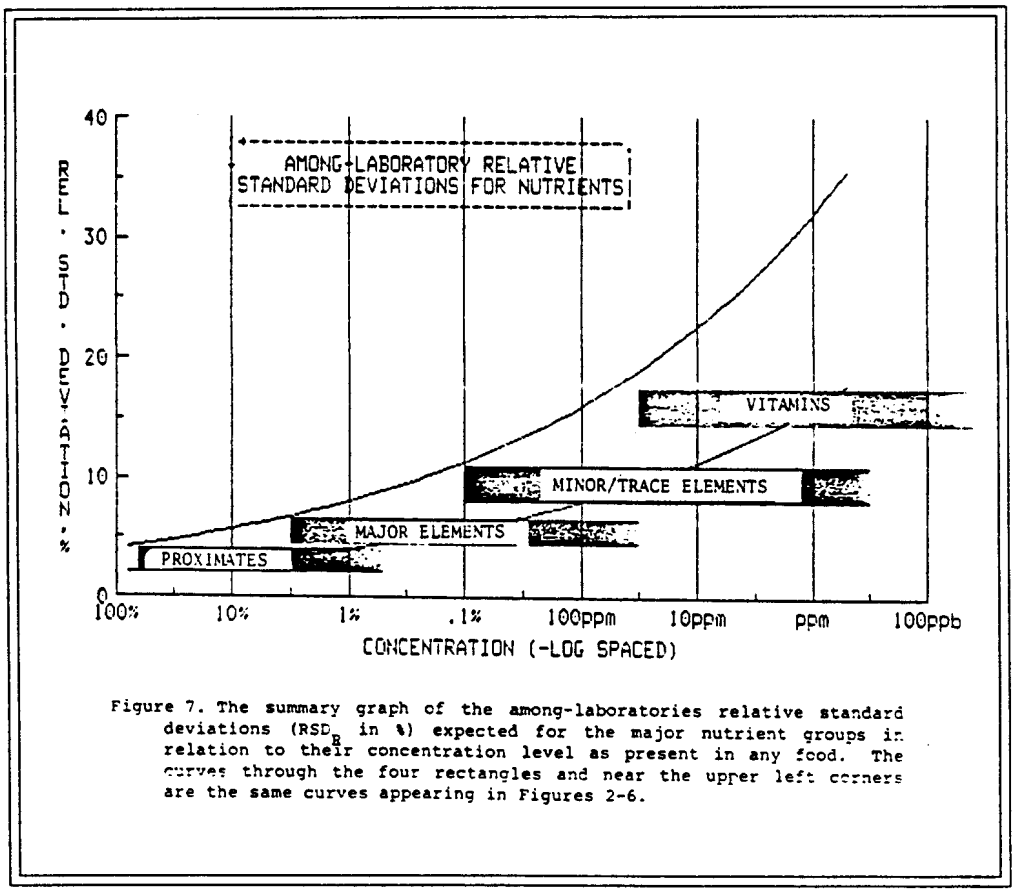

Figure 7. The summary graph of the among-laboratories relative standard deviations (RSD_R in %) expected for the major nutrient groups in relation to their concentration level as present in any food. The curves through the four rectangles and near the upper left corners are the same curves appearing in Figures 2-6.

tracting officers can relate similar experiences. In one case a request for proposals for pesticide residue analyses elicited interest from several dozen potential contractors. The number dropped to a half dozen when it was disclosed that the award would be made in part on the basis of results from actual blind analyses of the commodities of interest.

Conclusions

As a result of our review of the results of approximately 6000 interlaboratory studies conducted under fairly well-controlled conditions, we have constructed Figure 7, a summary of the RSD_Rs to be expected for the results of analyses for the various nutrients by a group of laboratories. The values do not include allowance for variability inherent in the commodity itself or from growing, processing, and sampling factors. (If these factors are also present, the standard deviations must be squared to obtain the variances, the variances should be added, and the square root of the sum should be taken to obtain the final "total" standard deviation. This standard deviation is transformed to an RSD by dividing by the mean and then multiplying by 100. A better value can be calculated with a scientific calculator (with an exponent key) from the general formula given at the bottom of Table 1, by inserting a specific concentration expressed as a decimal fraction.) The results are considered only as an approximate historical descriptive summary of the data of numerous studies. Any individual study can deviate considerably, by as much as a factor of 2 in either direction, and still be within the acceptable confidence interval.

The production of a database, however, requires the existence of absolute standards against which the values for specific analytes can be measured. In the case of proximate analysis, the analytical results are method-specific with no systemic error. No NIST standards existed during the data accumulation phases. Formulation for these types of analytes of reference materials that would remain stable over a reasonable period of laboratory storage would be difficult, although dried eggs and dried milks kept refrigerated and in a moisture- and oxygen-free atmosphere approach reference material requirements. But these foods are not useful for the important fiber-related analytes. Incorporation of a section in published papers describing the quality control efforts on an equal basis with the classical sections on materials, methods, results, discussion, conclusion, and references should become a necessary part of good manuscript preparation practices.

Table 2 - **Some typical and maximum among-laboratories acceptable relative standard deviations (RSD^R)** to be expected from historical analytical variability of nutritionally important analytes. Sampling and fabrication variability, if present, must be added vectorially (as variances).

Nutrient	Concentration			RSD_R	
	Mean	Unit	Range	Typical	Maximum
Proximates	10	%	100-0.5	2	5
Major Elements	0.1	%	5-0.005	5	10
Minor/Trace Elements	10	ppm	1000-0.5	10	20
Vitamins	1000	ppb	50-5000	15	35

References

(1)Food and Drug Administration (1990) *Code of Federal Regulations* Title 21, Section 2.19, U.S. Government Printing Office, Washington, DC 20402.

(2)Youden, W. J. (1967) *J. Assoc. Off. Anal. Chem.* 50, 1007-1013. *Statistical Manual of the AOAC,* Association of the Official Analytical Chemists, Arlington, VA 22201.

(3)Horwitz, W. (1988) *Pure Appl. Chem.* 60, 855-864; (1989) *J. Assoc. Off. Anal. Chem.* (1989) 72, 694-704.

(4)Peeler, J. T., Horwitz, W., and Albert, R. (1989) *J. Assoc. Off. Anal. Chem.* 72, 784-806.

(5)Horwitz, W., Albert, R., Deutsch, M. J., and Thompson, J. N. (1990) *J. Assoc. Off. Anal. Chem.* 73, 661-680.

(6)Horwitz, W., Albert, R., Deutsch, M. J., and Thompson, J. N. Under review for submission to *J. Assoc. Off. Anal. Chem.*

(7)Tanner, J. T. and Barnett, S. A. (1985) *J. Assoc. Off. Anal. Chem.* 68, 514-522; (1986) 69, 777-785; additional manuscript under review.

(8)Rasberry, S. (May 1991) *Am. Lab.* 23 (8), 59-60.

(9)Bigwood, D. W., Heller, S. R., Wolf, W. R., Schubert, A., and Holden, J. M. (1987) *Anal. Chim. Acta* 200, 411-419.

(10)Horwitz, W. (1978) *Anal. Chem.* 50(6), 521A-524A.

(11)Schubert, A., Holden, J. M., and Wolf, W. R. (1987) *J. Am. Diet. Assoc.* 87, 285-299.

(12)Lurie, D. G., Holden, J. M., Schubert, A., Wolf, W. R., and Miller-Ihli, N. J. (1989) *J. Food Comp. Anal.* 2, 298-316.

(13)Holden, J. M. and Davis, C.S. (1990) *Fresenius J. Anal. Chem.* 338, 476-478.