CRITICAL EVALUATION OF NUTRIENT DATA:

A PREREQUISITE FOR GENERATING NEW DATA

Joanne M. Holden
Nutrient Composition Laboratory
Beltsville Human Nutrition Research Center
Agricultural Research Service
U.S. Department of Agriculture
Beltsville, Maryland 20705

The research goal of the Nutrient Composition Laboratory (NCL), ARS, USDA is to provide essential data on the nutrient content of foods as consumed in the United States (U.S.). In order to make efficient use of limited fiscal and human resources, we must evaluate the quality of available data for a given nutrient. If high quality composition data exist for levels of a given nutrient in a wide variety of foods there is no need to generate more. If the quality of composition data has been determined to be inadequate, foods to be analyzed must be ranked in order of their priority for analysis. The NCL is particularly interested in the list of foods known as 'core foods,' that is, those foods most frequently consumed by the United States population. In particular, we have attempted to construct a list of 'core foods' for selenium (Se), defined as those frequently consumed foods which contribute the bulk of Se to the American diet. In addition we are interested in foods which may not be frequently consumed but are significant sources of Se.

Nutrient composition data are obtained from many sources - the food industry, published scientific literature, and from studies, published or unpublished, which are specifically designed to yield nutrient composition data. The job of evaluating these data from various sources for inclusion in the National Nutrient Data Bank is accomplished by the Nutrient Data Research Branch, Human Nutrition Information Service, USDA, under the direction of Frank Hepburn. In the past we have used their evaluation of such data as a guide to establishing analytical priorities within NCL for determining the nutrient composition of a specific food group or product type. Past projects have included fresh beef and pork, salty snacks, and fruit juices.

Recently, we have initiated a research project to establish guidelines for the evaluation of data for a specific nutrient - selenium (Se). The objective of this paper will be to describe the development of a system of general criteria to be specifically defined or qualified for this nutrient. In addition, I will briefly describe the method for selecting the list of 'core foods' which were the focus of the evaluation. Finally, I will show the application of the specific criteria requirements to published Se data to yield a list of 'core foods' for Se, together with a 'confidence code,' an indicator of quality for each mean value for each food.

As mentioned previously, the evaluation of available data is critical to the development of strategies for generating new data. The criteria which were developed and will be discussed were applied to published Se data. Selenium data were selected as a test case due to the current interest in the possible relationship between cancer incidence and Se intake (1,2). As a result of this interest, many human studies of Se intake and utilization are in progress and in need of Se composition data. In the 1980 edition of the Recommended Dietary Allowances a range of 50 to 200 ug/day has been given as the Estimated Safe and Adequate Daily Intake of Se for age 7 and over (3). An additional reason for focusing on Se is that good analytical methods for its determination are available. Finally, the data set is relatively small.

In order to evaluate published Se composition data for foods, two major tasks needed to be accomplished. First, we needed to develop a list of foods ranked in order of their importance to Se nutriture. The importance of a food to Se nutrition is based on its frequency of consumption, the amount consumed, and the Se concentration in that food. Frequency is defined as the rate or number of times a food is reported per day or per 3-day period. For example, if margarine was consumed at each of three meals then the frequency per day is 3. The list of foods was needed in order to establish priorities for consideration. A second but parallel activity concerned the development of a set of criteria to be used to evaluate the data from the scientific literature for the frequently consumed foods.

The frequency distribution for all foods consumed over a 3-day period by the 28,032 participants in the 1977-78 Nationwide Food Consumption Survey was obtained from the Food Monitoring Division, HNIS, USDA. This was considered to be the first step in an attempt to determine the contribution of any food to Se nutriture. We multiplied the frequency/day for each food by the average portion size to get the total intake/day of each food by the population. Frequency alone does not reflect the importance of a food's contribution since a food may be frequently consumed but may be consumed only in small quantities. Other foods may be consumed infrequently but their portion sizes may be relatively large. After evaluating the published data for various foods and determining a mean Se value for each, we incorporated that value into the frequency data for each of the top ranking foods to determine the Se contribution of each food to the U.S. diet.

Gram weights of some foods were summed to yield a total quantity of food aggregate consumed based on similarity of product type, the availability of Se data for each food, and the proximity of their Se levels.

Nationwide Food Consumption Survey food item	Total food intake/day
	g
Bread, whole wheat Bread, whole wheat, toasted Bread, whole wheat, high fiber Bread, whole wheat, high fiber, toasted	141,977

A food aggregate was regarded as a single item when the list of items was ranked by the total food intake per day. The amounts of the individual foods within the aggregate were retained in the file to permit re-grouping, if necessary.

An equally important and parallel activity was the development of the criteria for evaluating data. Five criteria were identified as significant in the evaluation of a data source: Analytical Method, Sample Handling, Sampling Plan, Analytical Quality Control, and the Number of Samples analyzed. These general criteria were similar to the criteria developed for the provisional Iron Table (5). In that publication three criteria were used. For each of the general criteria described here, a rating scale of '0' to '3' was developed. A range of 4 intervals was selected because it provided a level of resolution necessary for discriminating among various data sets and yet was not too detailed for available data. Requirements for each of the ratings were written to be specifically applicable to Se data. Writing the requirements required a knowledge of analytical methods for Se as well as a knowledge of sampling and statistics. In general a rating of '3' was considered to be optimum for a given criterion; a rating of '0' was assigned when there was no documentation or when the documentation was not acceptable. Table 1 lists the criteria and their specific requirements.

It is important to note some of the issues which are addressed by each of the criteria. For Analytical Method the emphasis is on method validation. The use of a validated official method or a well documented and validated new method is desirable. In order to obtain a rating of '3' the scientist must document the use of a validated method, including the use of tested quality control materials such as the National Bureau of Standards-Standard Reference Materials. Nutrient concentrations should be determined at quantifiable levels in the prepared sample extract or digest.

Sample Handling includes complete documentation of the validation of the homogenization method as well as details of sample preparation (cooked or uncooked, peeling, etc.) and storage. Evidence must be provided to indicate monitoring of the moisture content of samples during sample preparation and storage. Such precautions can assure that the sample which was analyzed is representative of the product which was purchased.

The criterion Sampling Plan addresses the representativeness of data for a specific nutrient and food aggregate. Our evaluation was conducted to determine the suitability of data for use in assessing the Se content of the American diet. With that purpose in mind, a study of the Se composition of samples obtained on a nation-wide basis with evidence of the representativeness of brands and varieties selected will be rated '3.' Minimal credit, '1', will be given to the references where samples were obtained only at the local level and where no evidence of the representativeness of brands or varieties is given.

Number of Samples Analyzed addresses the number of analyses for a given food in a particular study. Assuming that all other factors were equal, a study which reported multiple analyses of a food would be a stronger study

Table 1. Data Quality Criteria Requirements

Criteria\Ratings	3	2	1	0
Number of Samples	>10; SD, SE, or raw data reported	3-10	1-2; explicitly stated or not specified	-
Analytical Method	Official fluorometric (ref. given) or other method documented by a complete published write-up with validation studies for foods analyzed, including use of appropriate SRM where available, 95-105% recoveries on food similar to sample analyzed in same or other paper; Se concentration above quantitation limit of the method	Modified fluorometric or other method, some documentation, incomplete validation studies for foods analyzed; must include 90-110% recoveries on food similar to sample analyzed (or good recovery but no statistics given), and/or use of other method (official fluorometric, isotope dilution, or NAA) on same sample with good agreement (within 10%)	Non-fluorometric method, partially described; 80-90% or >110% recoveries on food similar to sample; or use of comparison method or recoveries on food only somewhat related to sample (animal/plant)	No documentation of method, no ref. or inaccesible ref. given, no validation studies, or poor agreement (>10%) of test method with comparison method on same sample
Sample Handling	Complete documentation of procedures incl. validation of homogenization method, details of food preparation and storage and moisture changes monitored	Pertinent procedures documented, seem reasonable, but some details not reported	Only edible portion analyzed	Totally inappropriate procedures or no documentation of criteria pertinent to food analyzed
Sampling Plan	Multiple geographical sampling with complete description; sample is representative of brands/varieties commonly consumed or commercially used	1 or 2 geographic areas sampled; sample is representative	Sample representative of small % of U.S. and/or origin not clear	Not described or sample not representative
Analytical Quality Control	Optimum accuracy and precision of method monitored and indicated explicitly by data	Documentation of assessment of both accuracy and precision of method; acceptable accuracy and precision	Some description of minimally acceptable accuracy and precision of method	No documentation of accuracy and/or precision

than one which reported single measurements. With multiple measurements (greater than 4-5) it is possible to generate some statistical indication of variability as well as a mean value.

Analytical Quality Control evaluates the level of accuracy and precision achieved in the day-to-day execution of the method. Both accuracy and precision are explicity defined in a paper to be published concerning the development of the criteria (6).

Greater detail on each of the criteria can be found in the two papers to be published within the next year (6,7).

The worksheet for whole wheat bread illustrates the assignment of ratings to each criterion and study. The reader should note that some references include data from more than one study. After all the data for a specific food have been rated, the numeric scores for each study are averaged over the five criteria to obtain a Quality Index (QI). For the Se data evaluation we decided that the presence of a '0' in Analytical Method for any study would automatically override the usual procedure for calculating the QI and yield a QI of zero for that study. The rationale was that if the analytical method was undocumented or not validated, the quality of data from that reference would be questionable. Similarly, ratings of zero in three criteria for a given study would yield a QI of zero.

WORKSHEET FOR WHOLE WHEAT BREAD

	Data Quality Criteria						Se	
Ref. No.	No. of Samples	Anal. Method	Sample Handling	Sampling Plan	Anal. Quality Control	Quality Index	Mean	SD
							µg/100g	
1	1	2	1	2	0	1.2 ↓	66.5	
2	1	2	0	2	0	1.0 ↓	53.5	
3	2	2	2	3	1	2.0 ↓	28.3	10.1
3	2	2	2	3	0	1.8 ↓	29.3	7.9
4	1	2	0	0	0	0	350	
5	1	2	1	2	0	1.2	67.0	
5	1	0	2	2	0	0	75 - 150	
						Σ = 6.0		
				Confidence Code = b		Grand Mean =	44.4	

↓ An index >1.0 is required for inclusion of an individual Se value in the calculation of the grand mean

After QIs were determined for each study, the sum of the QIs equal to or greater than 1.0 was calculated for each food. The worksheet for whole wheat bread indicates a QI sum of 6.0. The QI sum is used to determine a Confidence Code (CC) for the mean Se value for each food. The CC for a food indicates the degree of confidence a user can have in the mean value. The table below lists three ranges of QI's and their corresponding CC.

Assignment and Meaning of Confidence Codes

Sum of Quality Indices	Confidence Code	Meaning of Confidence Code
>6.0	a	Considerable confidence
3.4 to 6.0	b	Some confidence/ some problems
1.0 to <3.4	c	Less confidence/ limited data

Whole wheat bread received a CC of 'b' based on its QI sum of 6.0.

The mean Se concentration was calculated from those studies with QIs equal to or greater than 1.0. Those studies receiving a QI of less than 1.0 were excluded. None of the acceptable references was weighted more than another. The mean Se value for whole wheat bread was 44.4 ug/100 g.

The mean Se value for each food as determined by the evaluation procedure was combined with the total intake (g/day) of each food (as described above) to determine the Se intake/day for the population contributed by each food or food aggregate.

Determination of Se Intake/Day

Nationwide Food Consumption Survey food item	Total food intake/day	Se concentration	Se Intake/ day*
	g	ug/g	ug
Bread, whole wheat Bread, whole wheat, toasted Bread, whole wheat, high fiber Bread, whole wheat, high fiber, toasted	141,977	0.444	63,038

*Se intake/day by 28,030 survey respondents

This list of foods was ranked by the Se/day contribution to obtain the 'core foods' list specific for Se.

The evaluation of 65 references published since 1960 yielded mean minimum, and maximum Se values with Confidence Codes for more than 120 food items. The five highest ranked foods (beef, white bread, eggs, chicken, and pork) contribute one-half of the Se in American diets. Table 2 lists the first 16 foods which have been evaluated.

Based on this evaluation of available published data, NCL has begun to plan for further analyses of foods to obtain additional Se data. There were no Se data available for some foods in the NFCS list. Similarly, there were no data or frequency information for foods which have been introduced since the 1977-78 survey and may be important sources of Se. Select foods from these two groups were combined with high ranking foods which had been assigned a confidence code of 'c.' This group of foods was sampled and analyzed by NCL to supplement published data. In the future these data will be published and submitted to the Nutrient Data Bank, USDA.

Table 2. FIRST SIXTEEN SELENIUM CORE FOODS RANKED BY CONTRIBUTION TO U.S. DIETS

Rank	Food Item	Mean	Min. - Max.	Confidence Code	n/N+	Cumul. %¥
		µg Se/100g*				
1	Beef, ckd.	26	15 - 52	a	11/17	16.66%
2	White bread	32	23 - 54	a	7/9	31.05%
3	Eggs, ckd. (incl. scrambled, fried, soft-boiled)	25	16 - 38	a	7/11	39.02%
4	Pork/ham, fresh/cured, ckd./cnd. (incl. roasted, pan-ckd.)	35	19 - 92	a	8/13	46.46%
5	Chicken, ckd. (incl. fried, roasted)	21	17 - 26	a	5/10	52.90%
6	White rolls	34	21 - 61	a	4/4	57.00%
7	Whole milk	1.6	1.1- 2.5	b	4/15	60.20%
8	Whole wheat bread	44	28 - 67	a	4/7	63.25%
9	Tuna, cnd. (incl. light, white, packed in oil or water)	72	37 - 115	a	14/23	65.17%
10	Egg noodles, macaroni, spaghetti, ckd.++	20	14 - 42	a	20/23	67.03%
11	2% fat milk	2.6	---	c	1/3	68.88%
12	White rice, enriched, ckd.	9	---	c	1/5	70.59%
13	Macaroni & cheese, prep. from box mix	14	13 - 15	b	2/2	71.97%
14	Mayonnaise	60#	---	c	1/3	73.12%
15	Spaghetti with meat sauce, homemade	9	---	b	2/3	74.27%
16	Luncheon meat (incl. Spam¶)	28	27 - 28	c	2/5	75.42%

* Edible portion
\+ n = number of Se values from acceptable studies which were used to derive the mean value; N = total number of studies evaluated
¥ Cumulative % of total Se intake from core foods.
† Total Se intake/person/day from 89 core foods based on Nationwide Food Consumption Survey data and the mean Se values presented here = 74 µg
++ Se values calculated from dry and cooked values
\# Questionable based on the data from the other two studies which had acceptable QIs but reported semiquantitative Se levels ± 8 µg
¶ Geo. A. Hormel & Co., Austin, MN 55912

REFERENCES

1. Shamberger, R.J., Tytko, SA., and Willis, C.E.: Antioxidants and cancer. Part VI. Selenium and age-adjusted human cancer mortality. Arch Environ Health 31:231-5, 1976.

2. Schrauzer, G.N., White, D.A., and Schneider, C.J.: Cancer mortality correlation studies--III. Statistical associations with dietary selenium intakes. Bioinorg Chem 7:23-34, 1977.

3. Food and Nutrition Board, National Research Council: Recommended Dietary Allowances. 9th Rev. Ed. Washington, DC: National Academy of Sciences, 1980.

4. United States Department of Agriculture, Human Nutrition Information Service: Food Intakes; Individuals in 48 States Year 1977-78. Nationwide Food Consumption Survey 1977-78 Report No. I-1, 1983.

5. Exler, J. Iron Content of Food. Home Economics Research Report Number 45, Consumer Nutrition Division, Human Nutrition Information Service, USDA, Washington, DC, 1985.

6. Holden, J.M., Schubert, A., Wolf, W.R. and Beecher, G.R. A system for evaluating the quality of published nutrient data: Selenium, a test case. Food and Nutrition Bulletin, UN University 1986 (In press).

7. Schubert, A., Holden, J.M. and Wolf, W.R. Selenium content of a core group of foods based on a critical evaluation of published analytical data. Submitted to J. Amer. Diet. Assn.