USING RELIABILITY CODES

Joanne M. Holden, M.S.

Nutrient Composition Laboratory
Beltsville Human Nutrition Research Center
Agricultural Research Service
U.S. Department of Agriculture

The accuracy and precision of nutrient composition data have a significant impact upon the assessment of nutrient intake and the evaluation of relationships between intake and the incidence of certain disease states (1). Many users of nutrient data assume that a value which appears in print or in a data tape is of the highest quality and absolutely correct. However, the quality of an individual data point in a database or table is dependent upon several factors and merely represents an estimate made from available sources. For a single nutrient, the values for some foods may be accurate and precise; for other foods the values may be of poor quality due to the limitations of the source(s) from which the value was taken. Quality indicators of nutrient composition data quality could provide users of such data with information to determine the reliability of estimates of nutrient intake. In generic terms, such quality indicators could be called "reliability codes".

As mentioned above, a nutrient value for a specific food represents the compilation of values for that nutrient and food taken from various sources. Therefore, the quality of each tabular value is dependent upon the quality of those individual values. Furthermore, a published study may contain nutrient composition data which may be appropriate in its original context but may not be suitable for use in a nutrient data bank. For example, the primary objective of a study may be to develop analytical methodology and may include nutrient composition data as an example of methodological performance. For such a study, the selection of samples may be limited. Similarly, the objective of a study may be to test the effects of a new feeding regime on the level of a particular nutrient which results in the animal consuming that diet or to evaluate a new crop variety not currently marketed. Each of these situations would produce nonrepresentative data and would not be appropriate for use in assessments of nutrient intake by population groups.

In addition to the lack of representativeness of a particular value, the accuracy and precision of a nutrient value for a specific food taken from a single source is dependent upon several other factors. One factor concerns the appropriateness of the analytical method used, including its validation by reference materials or by another definitive method. Furthermore, the evidence of analytical quality control, satisfactory execution of the method on a day-to-day basis, is necessary for determining the quality of a single value. In addition, the sampling plan which had been used for selecting the samples should be fully documented and appropriate. Finally, sample handling techniques and the number of samples analyzed must be known in order to determine the soundness of the estimate.

A system for evaluating published data for these various factors or categories and for quantifying the degree to which each of these requirements has been met has been developed by the Nutrient Composition Laboratory (NCL) for the evaluation of selenium (Se) data (2,3) and, more recently, modified for the evaluation of copper data (4). For illustrative purposes, Table 1 contains a portion of the information presented in the article by Schubert et al. (3). This system provides detailed documentation of the data which have been compiled from various published sources to yield a single nutrient (e.g., Se) mean for a given food product. In addition to the mean selenium value, the table includes the references for acceptable studies, that is, studies whose mean values have been included in the computation of the grand mean. Also, the numbers

Table 1. Selenium content of selected foods*

food aggregate	mean	min.	max.	no. of Se values accept.	no. of Se values total	Conf. Code	ref. no. of acceptable studies
	μg Se/100g						
BEEF, LAMB, PORK, VEAL							
Beef, raw	22	5	42	19	27	a	10,12,14,15,18-24
Beef, ckd.	26	15	52	11	17	a	9,12,25
Beef gravy	1.0			1	2	c	12
Beef liver, raw	40	18	63	7	10	a	10,12,14,15,18,19,26
Beef liver, ckd.	56	43	71	3	5	b	9,12
Lamb, raw	21	6	32	7	11	a	10,12,15,18,19,27
Lamb, ckd.	17			1	5	c	9
Heat loaf, beef, ckd.	17	12	27	3	3	b	9,12
Pork/ham, fresh/ cured, raw	33	19	51	8	12	a	10,12,14,15,19,28
Pork/ham, fresh/cured, ckd./cnd. (incl. roasted, pan-ckd.)	35	19	92	8	12	a	9,12,29
Veal, raw	28	20	35	2	2	c	15
Veal, ckd.	12	12.0	12.3	2	2	b	9

*Taken from Schubert et al. (3)

of acceptable studies and of studies evaluated are provided. In addition, the range of acceptable values is given. Finally, a confidence code ("a", "b", or "c"), indicating the degree of confidence a user of nutrient composition data can have in a particular value, is provided. For example, a confidence code of "a" indicates the user can have considerable confidence in the mean. In contrast, a code of "c" indicates the user can have less confidence in this value due to the limited quantity and/or quality of data.

Data which have been evaluated in a systematic way and coded to indicate the quality of the individual means may be used in several ways. Initially, Se data were evaluated by the NCL to determine where resources for additional analyses should be used. Foods were ranked by their contribution to the population's daily intake of this nutrient. Foods which had "c" codes or no acceptable data were designated for further analysis. High ranking foods were included in the list for verification analysis regardless of their code. Finally, new foods which did not appear in the original list were considered for limited analysis to obtain "ballpark values". As additional analyses were completed the individual data were added to the previously evaluated data for that food; new means and confidence codes were generated (5). As one can see, the process is iterative. In fact, the process of improving the Se composition data is a three year effort. Following a pilot study of eighty foods collected in each of three cities, a more extensive sampling in five to nine cities is currently in progress. Beginning in late 1987, a second sampling of five to nine cities will be conducted to provide more data and to investigate more fully the variability of selenium in foods.

The published systematic evaluation of nutrient data can provide several different kinds of information to users of food composition data. In addition to the mean for a single food, the inclusion of confidence codes provides the user with some measure of the quality of each mean. Means and confidence codes for foods can be grouped by food category, e.g. red meats, breads, fruits and vegetables, etc. The user can view the distribution of confidence codes within a food category. For example, in Table 1, the range of confidence codes for Se values within the Beef-Pork-Veal-Lamb category indicates six "a" codes, three "b" codes, six "c" codes (3). This is an important and, perhaps, comforting observation since beef and pork are ranked as the first and third contributors of Se as determined from the data evaluation. In contrast, the tally for fruits and vegetables indicates that five codes were "c", four were "b", and three were "a". In view of the fact that fruits and vegetables do not provide a significant amount of Se to the diet of the population, the poorer quality of Se data for this category is not so disturbing.

In the evaluation of Se data (3), data for the first 10 ranked food items were of high quality (7 "a's" and 3 "b's"). Beyond the tenth food item the three confidence codes appear in random order, with "b's" and "c's" approximately equally represented and "a's" occurring 30% as frequently as either "b" or "c". If these data were used to calculate the Se contents of diets for a population, the user could have good confidence in the estimates of population intake obtained. Those frequently consumed foods which are important contributors of selenium are characterized by good data. In general, less frequently consumed foods with lower selenium concentrations have data of poorer quality. In view of this, estimates of intake would be fairly reliable. In the future it would be possible to tally the confidence codes for a specific nutrient for all food items consumed in a survey to determine the overall quality of data for foods actually consumed. Foods of particular concern could be designated for additional analyses.

In addition to seeing the quality indicated by the confidence code, the user of data can see the list of individual sources which were evaluated and compiled to obtain a mean value. A concerned user could review a specific reference. Furthermore, compilers of nutrient data could determine if the nutrient mean which appears in the data base or tape includes values from an additional reference or source at hand. In this way the user could avoid duplicate entry and double weighting of a given value.

Finally, a well-documented system of evaluating data provides generators of nutrient data with guidelines which can be incorporated into planning for new studies. Similarly, editors and reviewers can use guidelines to facilitate the review process. In fact, the various criteria which have developed for the categories in the Se system have been reviewed and modified by the editors of a new journal, Journal of Food Composition and Analysis, to serve as guides to authors. This journal will be published, starting in 1988.

More than ever, users of food composition data are requiring accurate and precise nutrient composition data to evaluate the relationships between diet and health. Those of us who generate data, manage databases, or facilitate the flow of data in some other way must provide quality indicators or some other documentation as assurances of data quality and reliability.

REFERENCES

1. Food and Nutrition Board, National Research Council: Nutrient Adequacy, Assessment Using Food Consumption Surveys. National Academy Press, Washington, D.C. 1986.
2. 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 (Supplement), United Nations University, 1987. (In press).
3. Schubert, A., Holden, J.M., and Wolf, W.R. Selenium content of a core group of foods based on a critical evaluation of published data. Journal of the American Dietetic Association. 87 (3):285, 1987.
4. Schubert, A., Lurie, D.G., Wolf, W.R. and Holden, J.M. The copper content of foods: A critical evaluation of published analytical data. 70th Annual Meeting of the American Dietetic Association, October 19-23, 1987. (Abstract submitted).
5. Lurie, D.G., Schubert, A., Holden, J.M., and Wolf, W.R. Assessment and improvement of selenium composition data. 19th Annual Meeting of the Society of Nutrition Education, July 6-9, 1986. (abstract).