

DEVELOPMENT OF THE CONTINUING SURVEY OF FOOD INTAKES BY INDIVIDUALS (CSFII) DATA BASE

Betty P. Perloff

The USDA Nutrient Data Bases for Individual Food Intake Surveys are a series of nutrient data bases developed for use in the USDA Nationwide Food Consumption Surveys. The first release (1) of this series was used to process data for the Nationwide Food Consumption Survey conducted in 1977-78. The second release (2) was prepared for the Continuing Survey of Food Intakes by Individuals (CSFII). Many differences in food descriptions and food codes exist between the two releases. Also, Release 2 contains updated nutrient values as well as data for several additional nutrients.

The Continuing Survey of Food Intake by Individuals is a nationwide survey in which 24-hour interviewer-assisted recalls are taken from survey respondents approximately every two months. CSFII 85 was conducted between April 1985 and April 1986. CSFII 86 began in April 1986 and ends in April 1987. For each year, national samples were drawn for women 19 to 50 years of age and their children 1 to 5. A low income sample was included for the women and children. CSFII 85 also included a sample of men 19 to 50 years of age.

Before CSFII 85 began, a nutrient data base containing data for approximately 4,500 items was prepared. As the survey progressed, data for new foods were added as needed. This nutrient data base was released to the public in May 1986 as Release 2 of the nutrient data base series for individual food intake surveys. Approximately 3 months of CSFII dietary records had been coded. The nutrient data base was released again as Release 2.1 (3) of the series after coding was completed for CSFII 85. Release 2.1 is identical to Release 2 except that it contains data for about 500 additional food items.

Release 2 was also used by the National Center for Health Statistics, U.S. Department of Health and Human Services, to process data collected during the Hispanic Health and Nutrition Examination Survey (NHANES) for the Mexican-American population in the southwestern United States.

The same nutrient data base that was used for CSFII 85 is being used for CSFII 86, but as coding of dietary records progresses, new foods are added as needed. When the food consumption data are released for CSFII 86, an updated version of the nutrient data base will also be released.

The remainder of this paper discusses the CSFII nutrient data base and its development and describes its relationship to two other nutrient data bases prepared by the Nutrition Monitoring Division (NMD) at the USDA's Human Nutrition Information Service (HNIS).

Food codes

Food codes are 7-position numeric codes. The first position represents a major food group; positions 2-3 represent subgroups. For example, all foods with "2" in the first position are meat items, those with "21" in positions 1-2 are beef items, and those with "211" in positions 1-3 are beef steak items. The nine major food groups are listed in Table 1.

Table 1. Major Food Groups for CSFII

1. Milk and milk products
 2. Meat, poultry, fish, and their mixtures
 3. Eggs and their mixtures and substitutes
 4. Dry beans, peas, other legumes, nuts, and seeds
 5. Grain products
 6. Fruits
 7. Vegetables
 8. Fats, oils, and salad dressings
 9. Sugars, sweets, and beverages
-

Foods

Foods are described in as much detail as necessary to distinguish between different forms having different nutrient profiles. For example, eight items are present for spinach:

1. spinach, fat added in cooking
2. spinach, fat not added in cooking
3. spinach, not specified as to fat added in cooking
4. spinach, raw
5. spinach, creamed
6. spinach souffle
7. spinach with cheese sauce
8. spinach and cheese casserole

Food items containing nonspecific or general descriptions were assigned nutrient values for a commonly eaten form of the food or for a composite of more than one form. For example, the third spinach item above was assigned nutrient values for "spinach, fat added in cooking," but nutrients for the item "ground beef, not specified" were developed by combining data for regular, lean, and extra lean ground beef. Items with non-specific descriptions were included because experience has shown that survey respondents cannot always be specific about the foods they have eaten.

Development of nutrient profiles

The nutrients for which data are included on the three nutrient data files discussed in this paper are listed in Table 2. In selecting the food components to include in the CSFII nutrient data base, the key consideration was the needs of the scientific community, but adequacy of available data, by necessity, was taken into account. The final selection of 28 food components plus food energy was made jointly between HNIS and the National Center for Health Statistics.

Since nutrient data are not available for most mixed dish items, a procedure to calculate nutrient values based on recipes was used extensively. General nutrient retention factors were applied to account for cooking losses. All nutrient values used in the development of the CSFII data base were first drawn together in the Primary Nutrient Data Set for USDA Food Consumption Surveys (PDS), which is described below. Foods on the PDS include not only commonly consumed items but also those normally used only as ingredients in mixed dishes such as baking powder.

DEVELOPMENT OF THE CSFII DATA BASE

Table 2. Nutrient Data Bases
Development of the Continuing Survey of Food Intakes by Individuals (CSFII) Data Base, by Betty Perloff

	Individual Food Intake Survey (for CSFII)	Primary Nutrient Data Set for USDA Food Consumption Surveys	Standard Reference	
			Updated	Not Updated
<u>Proximates:</u>				
Water	x	x	x	x
Energy	x	x	x	x
Protein	x	x	x	x
Total lipid	x	x	x	x
Total carbohydrate	x	x	x	x
Crude fiber			x	x
Total dietary fiber	x	x	1	
Ash			x	x
<u>Minerals:</u>				
Calcium	x	x	x	x
Iron	x	x	x	x
Magnesium	x	x	x	
Phosphorus	x	x	x	x
Potassium	x	x	x	x
Sodium	x	x	x	x
Zinc	x	x	x	
Copper	x	x	2	
Manganese			2	
<u>Vitamins:</u>				
Ascorbic acid	x	x	x	x
Thiamin	x	x	x	x
Riboflavin	x	x	x	x
Niacin	x	x	x	x
Pantothenic acid			x	
Vitamin B-6	x	x	x	
Folacin	x	x	x	
Vitamin B-12	x	x	x	
Vitamin A (IU)	x	x	x	x
Vitamin A (RE)	x	x	x	
Carotene (RE)	x	x		
Vitamin E	x	x	3	
<u>Fatty Acids:</u>				
Total saturated	x	x	x	x
Total monounsaturated	x	x	x	
Total polyunsaturated	x	x	x	
Oleic			x	x
Linoleic			x	x
Others			x	
Cholesterol	x	x	x	x
18 amino acids			x	
Alcohol	x	x		

- 1 Only a limited number of values are present for neutral detergent fiber
- 2 Values are not included for dairy and egg products or spices and herbs
- 3 Only a limited number of values are present

Table 3. CSFII Recipe Examples

<u>CSFII Item</u>		<u>PDS Items</u>		
Number	Name	Number	Name	Grams
1. 611-1901	Orange	09200	Orange	100
2. 622-0101	Ambrosia	09200	Orange	378
		09040	Banana	357
		12104	Coconut	120
		92310	Confectioner's Sugar	30

Table 4. CSFII Recipe Fats

1. Olive oil
2. Corn oil
3. Soybean oil
4. Soft margarine, unspecified oils
5. Regular margarine, unspecified oils
6. Imitation margarine (approximately 40% fat), unspecified oils
7. Margarine-like spread (approximately 60% fat), unspecified oils
8. Butter
9. Lard
10. Shortening, hydrogenated soybean-cottonseed oils

A computerized file was constructed to link each item on the CSFII data base to the PDS. This file also links CSFII items that are mixed dishes to all appropriate ingredient items on the PDS. This linking file is actually a file of recipes, with items that are not mixed dishes coded as single-ingredient recipes. Examples of single and multi-ingredient recipes are presented in Table 3. The CSFII data base was generated by a computer program using the recipe file to determine which nutrient values on the PDS to use for each CSFII item. The program calculated the nutrient content of mixed dishes as necessary. The recipe calculation procedure and recipe file were discussed in detail at the Tenth National Nutrient Data Bank Conference (4).

Recipes containing salt as an ingredient were calculated both with and without the salt, and both sets of nutrient values appear on the data base. A special field in each record indicates if a set of values was calculated directly from the recipe or calculated by omitting salt from the recipe. Recipes including fat as an ingredient or recipes involving the absorption of fat during cooking were calculated in several ways--by using data for the type of fat specified in the recipe and also by substituting data for several other types of commonly used fats (Table 4). For example, if a recipe normally used butter as the ingredient, the nutritive values were calculated by using butter as the ingredient and also by using the other nine fats listed in Table 4. Complete sets of nutrient values for these different calculations are included in the CSFII data base. The type of fat used for each calculation is designated in a field in the data record.

DEVELOPMENT OF THE CSFII DATA BASE

Primary Nutrient Data Set for USDA Nationwide Food Consumption Surveys

The Primary Nutrient Data Set (PDS) was developed for use in generating and updating the USDA Nutrient Data Base for Individual Food Intake Surveys. It is continually updated as needed for survey purposes. The one and, to date, only public release of this data set was made in May 1986. It is available on magnetic tape with two other data sets used to create Release 2 of the USDA Nutrient Data Base for Individual Food Intake Surveys--the recipe file and the nutrient retention factors file (5). Together, they serve as documentation for that specific release of the survey nutrient data base. The next release probably will be made after CSFII-86 is completed and at the same time that a new USDA Nutrient Data Base for Individual Food Intake Surveys is issued.

Most of the data on this version of the PDS are from Release 5 of the USDA Nutrient Data Base for Standard Reference (described below) (6). Values were added for the nutrients missing from the Standard Reference Data Base, and complete nutrient profiles were added for missing food items. If analytical data were not available, the values were imputed from other forms of the foods or were estimated from data for similar foods. The values are for 100 grams of the edible portion of a food. Included with each value is a code to indicate whether or not it is from the Standard Reference Data Base and whether it is based on analytical data or is an imputed value. A date is included with each value not from the Standard Reference Data Base to indicate when it was added to this data set.

All items from the Standard Reference Data Base carry Standard Reference identification numbers. Added food items have been assigned special identification numbers. Values in the data base for carotene are those assumed by HNIS in arriving at the values for total vitamin A and should not be interpreted as representing solely beta-carotene. Values for beta-carotene content of foods have not been reported frequently, and existing reports are often not clear as to whether a value is explicitly for beta-carotene or whether it includes other carotenoids. Only limited analytical data are available for vitamin E and dietary fiber.

HNIS' working version of the PDS has been updated with new data from Release 6 of the Standard Reference data base (7) and is currently being used to develop the nutrient data base that will be used with the Nationwide Food Consumption Survey 1987.

USDA Nutrient Data Base for Standard Reference

This nutrient data base corresponds to Agriculture Handbook No. 8 "Composition of Foods . . . Raw, Processed, Prepared". It contains all the data that have been published in the revised sections of the handbook (8), and for those food groups for which revision is incomplete, it includes data from the 1963 edition (9). However, values for enriched flour and bread and for other products made with enriched flour have been changed to reflect the current revised standards of identity. The current release is the sixth for this data base. As new sections of the Agriculture Handbook 8 revision are published, new releases of this data base will follow.

As indicated in Table 2, data for more nutrients are present for the items taken from the revised sections of the handbook than are present for the items from the 1963 edition of the handbook. Where blank spaces appear in the handbooks, values were imputed and included in the data base, if possible. Nutrient values are present for 100 grams of the edible portions of each food and, for most items, two common household measures and 1 pound as purchased. Five-position numeric food codes are used.

This data set is available on both magnetic tape and floppy disk.

Summary

The current release of the USDA Nutrient Data Base for Individual Food Intake Surveys was developed for the Continuing Survey of Food Intake of Individuals 1985. All nutrient values used

in its generation are found on the Primary Nutrient Data Set for USDA Food Consumption Surveys. A computerized file links the two data sets and includes recipe ingredient information for calculating nutrient content of CSFII mixed dish items based on nutrient content of their ingredients. The main source of nutrient data on the PDS is the USDA Nutrient Data Base for Standard Reference.

REFERENCES

1. U.S. Department of Agriculture, Human Nutrition Information Service. 1980. USDA Nutrient Data Base for Individual Food Intake Surveys, Rel. 1. Springfield, VA: National Technical Information Service. Accession No. PB82-138504/HBF. Computer tape.
2. U.S. Department of Agriculture, Human Nutrition Information Service. 1986. USDA Nutrient Data Base for Individual Food Intake Surveys, Rel. 2. Springfield, VA: National Technical Information Service. Accession No. PB86-206299/HBF. Computer tape.
3. U.S. Department of Agriculture, Human Nutrition Information Service. 1987. USDA Nutrient Data Base for Individual Food Intake Surveys, Rel. 2.1. Springfield, VA: National Technical Information Service. Computer tape.
4. Perloff, B.P. Recipe Calculations for NFCS Data Base. In Proceedings of Tenth National Nutrient Data Bank Conference, July 22-24, 1985, ed. S. Murphy and D. Rauchwarter, pp. 221. Springfield, VA: National Technical Information Service. Accession No. PB86-159589.
5. U.S. Department of Agriculture, Human Nutrition Information Service. 1986. Data Sets Used to Create Release 2 of USDA Nutrient Data Base for Individual Food Intake Surveys. Springfield, VA: National Technical Information Service. Accession No. PB86-206281/HBF. Computer tape.
6. U.S. Department of Agriculture, Human Nutrition Information Service. 1985. USDA Nutrient Data Base for Standard Reference, Rel. 5. Springfield, VA: National Technical Information Service. Accession No. PB85-140051/HBF. Computer tape.
7. U.S. Department of Agriculture, Human Nutrition Information Service. 1987. USDA Nutrient Data Base for Standard Reference, Rel. 6. Springfield, VA: National Technical Information Service. Computer tape.
8. U.S. Department of Agriculture. 1976- . Composition of Foods: Raw, Processed, Prepared. Agric. Handb. No. 8: AH-8-1, Dairy and Egg Products, 1976; AH-8-2, Spices and Herbs, 1977; AH-8-3, Baby Foods, 1978; AH-8-4, Fats and Oils, 1979; AH-8-5, Poultry Products, 1979; AH-8-6, Soups, Sauces, and Gravies, 1980; AH-8-7, Sausages and Luncheon Meats, 1980; AH-8-8, Breakfast Cereals, 1982; AH-8-9, Fruits and Fruit Juices, 1982; AH-8-10, Pork and Pork Products, 1983; AH-8-11, Vegetables and Vegetable Products, 1984; AH-8-12, Nut and Seed Products, 1984; AH-8-13, Beef Products, 1986; AH-8-14, Beverages, 1986; AH-8-16, Legumes and Legume Products, 1986.
9. Watt, B.K., and Merrill, A.L., 1963. Composition of Foods . . . Raw, Processed, Prepared. U.S. Dept. of Agriculture, Agric. Handb. No. 8, 190 pp.