REPORT FROM USDA'S NUTRIENT DATA RESEARCH BRANCH

F. N. Hepburn
Human Nutrition Information Service
U.S. Department of Agriculture

The Nutrient Data Research Branch of the Human Nutrition Information Service (HNIS) is responsible for compiling and making available information on the nutrient composition of foods. This work includes not only the publication of reference tables but also the creation of nutrient data bases used for interpreting results of food consumption surveys conducted by another branch of HNIS.

This presentation reports on the status of publications, describes the derivation of the Primary Nutrient Data Set recently created for the Continuing Survey of Food Intakes by Individuals, and briefly outlines activities that will generate new analytical data and improve the reliability of nutrient composition data.

STATUS OF PUBLICATIONS

The Nutritive Value of Foods, Home and Garden Bulletin No. 72 (HG-72), has been completely revised as to food items, nutrients listed, and nutrient values. New foods have been added and foods no longer on the market or no longer popular have been deleted, resulting in a total of 910 items in place of 730 in the previous edition.

Data from published sections of revised Agriculture Handbook No. 8 (AH-8) have been incorporated, and all data have been reviewed, compared to new analytical data, and updated as necessary to reflect current knowledge. For example, new data for iron in beef have been included in accordance with the information on hand for the revision of the AH-8 section on beef.

Responding to popular demand, we have added data on sodium and cholesterol. Vitamin A is given in terms of both International Units (IU), for the convenience of those who continue to work in the traditional units, and Retinol Equivalents (RE), so that reference can be made to the current Recommended Dietary Allowances and foods can be compared in terms of relative vitamin A activity. Fatty acid content is expressed as total saturated, total monounsaturated, and total polyunsaturated fatty acids.

HG-72 is intended primarily for the general public and for educators and health professionals who help American understand the nutrient content of foods they eat. We expect great interest in its use as a data base for microcomputers. The National Technical Information Service will have available both computer tapes and floppy discs of HG-72 as soon as possible after its publication.

The revisions of Agriculture Handbook No. 8 (AH-8) have not progressed as well as predicted at last year's conference. The following table shows current estimated completion dates:

Publication Schedule for Revision of AH-8

Section No.	Food Group	Completion Date
13	Beef products	Winter 1985
14	Beverages	Winter 1985
15	Fish and shellfish	Spring 1986
16	Legumes	Summer 1986
17	Lamb, veal, and game	Winter 1986
18	Bakery products	Summer 1987
19	Sugars and sweets	Winter 1987
20	Cereal grains, flours, and pastas	Summer 1987
21	Fast foods	Fall 1986
22	Mixed dishes	Fall 1987
23	Miscellaneous foods	Fall 1987

The section on beef has first priority. It is entering final computer stages at this time and we are expecting its appearance in printed form sometime during the winter of 1985. Because the section on lamb, veal, and game is being prepared by the same staff now working on beef products, publication of section 17 will depend upon the actual completion date of section 13.

Sections 18 through 23 require later scheduling because they will include results of analyses now being performed under contract. Bakery products and mixed dishes are expected to require longer preparation time because of the complicated nature of the groups. Section 19 is being prepared by the staff working on 14; section 22 by those working on 21.

DATA BASE FOR CONTINUING SURVEY OF FOOD INTAKE BY INDIVIDUALS

A major task this past year was establishing the data base for the new continuing survey. Two factors have made it time consuming. One was the decision to expand the coverage of nutrients from the 14 used in the 1977-78 Nationwide Food Consumption Survey (NFCS) to double that number. The added components are --

Copper	Alpha tocopherol
Potassium	Vitamin A (RE)
Sodium	Carotene (RE)
Zinc	Folacin
Alcohol	Saturated fatty acids
Cholesterol	Monounsaturated fatty acids
Dietary fiber	Polyunsaturated fatty acids

The data base is less reliable for some of these components and interpretation of results will have to consider this limitation. The relative degree of reliability of the various components is discussed below.

The decision to track sodium and the different groups of fatty acids made the task difficult because of the need to provide for foods both with and without salt and for foods prepared with different sources of fat.

The new data base required a new procedure, which is described in the paper by Betty Perloff. Essentially, the procedure involves linking the foods reported by survey respondents to a Primary Nutrient Data Set for Food Consumption Surveys (PDS) through a recipe linking file. The PDS consists of foods that either are consumed as individual items or that are constituents of a food described by a recipe. To the extent possible, data for items are from the USDA Nutrient Data Base for Standard Reference. Additional items are added as required to describe foods being reported by the survey.

It is convenient to regard the PDS as a supplemented Standard Reference Data Base, but this is not strictly true because there are foods in the latter that are not used in the PDS. Values for all nutrients must be provided for new foods and data for those nutrients not included in the Standard Reference Data Base must be supplied as well. The sources of data used for creating the PDS were documented by code according to the following scheme:

Code	Description of Data Source
1	Analytical data, or values calculated directly from analytical data, from revised AH-8 sections.
2	New provisional analytical data from Nutrient Data Bank for food group sections not yet revised.
3	Unrevised analytical data from computer data set 456-3 (based on the 1963 AH-8).
4	Imputed values from USDA Nutrient Data Base for Standard Reference.
5	Nutrient label claims (Breakfast cereals only; AH-8-8).
6	New imputed values for Primary Nutrient Data Set.
7	Assumed zero values.

Various meanings of the word "imputed" are described in Linda Posati's paper on this subject. I am using the word in the same sense as it is used in regard to AH-8 and the Standard Reference Tape. Values in the printed publication are analytical or are calculated in a very direct manner from analytical data; they

are not considered to be imputed. Missing values are left blank. For the Standard Reference Data Base, blanks are filled with imputed values and flagged to differentiate them from other values. These imputed values are our best estimates, usually based on values for a similar food or another form of the same food.

Source codes are attached to each nutrient value in the Primary Data Set. We can examine the data for each nutrient for reliability in terms of data source; that is, the extent to which they are based on analytical as opposed to imputed values. To determine the proportion of analytical to imputed data in the Primary Data Set, data from source codes 5 and 7 were first eliminated. Data from source

codes 1, 2, and 3 were combined as analytical values and those from source codes 4 and 6 were combined as imputed values. About 2,250 total values were found for each nutrient. At least 2,100 remained after dropping code 5 and 7 data except for vitamin B_{12}, carotene, and dietary fiber. These exceptions were due, of course, to assumption of the absence of these components in large numbers of foods. The proportion of analytical data was calculated as the percentage of total data coded as either analytical or imputed. Results are shown in the following table:

Percent of Data from Analytical Sources in Primary Data Set

90 percent or more analytical		Less than 90 percent analytical		
Component	All Foods	Component	All Foods	Major Sources
Calcium	97	Vitamin C	83	92
Protein	97	Vitamin A (IU)	80	89
Fat	96	Magnesium	75	72
Thiamin	91	Zinc	73	79
Riboflavin	91	Copper	67	71
Niacin	91	Vitamin B_6	64	72
Sodium	90	Vitamin B_{12}	64	70
Potassium	90	Vitamin A (RE)	61	73
Phosphorus	90	Folate	56	69
Iron	90	Carotene (RE)	54	88
		Dietary fiber	29	40
		α-tocopherol	28	39

It is evident that the proportion of analytical data is high for the more familiar nutrients that have been tracked over a longer period, equaling or exceeding 90 percent, whereas analytical data of the newest components are below 30 of percent data from all foods. The column on the far right of the table shows results for only major food sources of nutrients. For this analysis, we excluded those food sources that provide less than 5 percent of the U.S. RDA per serving portion for most of the nutrients. For other components, cut off values were vitamin A, 50 IU or 12.5 RE; carotene, 12.5 RE; α-tocopherol, 1.0 mg; copper, 0.1 mg; and dietary fiber, 2.5 g. Portion sizes were based primarily on the default amounts of the NFCS 1977-78, amended for some foods for improved uniformity. Except for magnesium, percent analytical of the better sources is greater than that for all foods, indicating that more of the imputed data are from relatively poor nutrient sources and that more analytical values are available for better sources. The magnesium exception reflects a more general distribution of that element in the food supply. Although the better sources of vitamin E and dietary fiber showed higher percentages of analytical data than for all foods, the values remain relatively low, indicating that our knowledge of these components is much weaker than our knowledge of other food components being studied.

It must be remembered that we have not considered in these calculations the relative quality of analytical data. For example, even though data for the better sources of carotene are 88 percent analytical, these data are based primarily on the AOAC procedures, which generally report total carotenes.

Although amounts may be reported in terms of beta carotene, they may not be specific for beta carotene itself. We expect that results from ongoing contracts supported by the Human Nutrition Information Service and research now going on in the Nutrient Composition Laboratory will provide new data on the separate carotenoid components.

REVISION OF THE NUTRIENT DATA BANK SYSTEM

Under a HNIS sponsored contract, work is nearing completion to revise the Nutrient Data Bank computer system to take advantage of advances in computer technology since the original system was developed ten years ago. The overall effect will be to improve the efficiency of operation. The major change for our staff is that the system will be interactive with our food group specialists. Each specialist will have access to his or her own data sets through a terminal and will be able to examine and compare individual data entries in the process of summarization.

Provision has been made for the future encoding of quality indices of the data so that we will be able to generate confidence codes that apply to the summarized values. This feature is now only in the planning stage, but it will become possible as a result of the revision.

EXTRAMURAL RESEARCH

HNIS funds food composition research in several universities and elsewhere. Most of these research contracts generate data on the nutrient content of foods to provide information needed for the revision of specific food-group sections of the handbook. A few, such as one on carotenoids, are component specific. One project of particular interest to the Nutrient Data Bank Conference, is the development of a Guide for Nutrient Data Users, the product of a cooperative agreement with Grace Petot of Case Western Reserve University. We plan to make this new guide available with the purchase of nutrient data tapes as well as by separate purchase.

New contracts have just been awarded on miscellaneous foods, frozen prepared dishes, and fried foods. These new contracts are the last of those intended for remaining AH-8 sections. Other major research activities will be carried out in cooperation with the Nutrient Composition Laboratory at Beltsville, MD. These will include studies on selenium in foods and a study on the nutrient composition of cookies and sweet-snack bakery items.

FUTURE ACTIVITIES

Efforts will continue on revising AH-8. In addition to following the schedule on remaining sections, work will begin on updating food group sections previously published.

The nutrient data base for the CSFII will be expanded as the survey continues. At the same time we will develop the data base for the household segment of the 1987 NFCS.

We appreciate the opportunity to communicate with data users at the annual Nutrient Data Bank Conference. Your input is especially valuable in providing direction for our efforts.