

USING USDA CONTINUING SURVEY DATA

Howard A. Riddick, Ph.D.

The Continuing Survey of Food Intakes by Individuals (CSFII) is a timely and unique source of information on the dietary intakes of women and children. Potential uses of this dietary data include comparisons of food and nutrient intakes of different population groups; comparisons of food intake behavior with other health-related behavior like smoking and exercise; and evaluations of diurnal, seasonal, and yearly variations in food intake. In addition to analytical work, the nutrient data tapes and related material used in processing the survey data are useful to people who conduct their own dietary intake surveys.

The core of the CSFII is a national sample of women 19 to 50 years of age and their children 1 to 5 years of age in the 45 conterminous states. This sample, sometimes referred to as the "core monitoring group", was selected because previous surveys have shown that women and young children are more likely than other population groups to have diets low in certain nutrients. The 1985 survey also includes a sample of low-income women and children in the same age ranges as the core monitoring group and a sample of adult men 19 to 50 years of age. The 1986 survey included newly drawn samples for both core and low-income women and children.

The CSFII contains many of the basic features of the individual intake component of the Nationwide Food Consumption Survey of 1977-78 (NFCS 1977-78). There are some differences, however. Information in NFCS 1977-78 was collected for 3 successive days using a 1-day dietary recall followed by a 2-day food record completed by respondents. CSFII data was collected using 1-day dietary recalls only. Men were surveyed once, while women and children from the core and low-income samples were surveyed on 6 separate days over a 1-year period. The first day of intake for CSFII was collected using a personal interview in the respondent's home. Subsequent days of data were collected by telephone at approximately 2-month intervals. Individuals in households without telephones were contacted in person.

Since the 1977-78 survey, the food codes, food code descriptions, quantity measures, and nutritive values have been revised. Revisions include a greater number and variety of products (such as low-sodium products) and updated information on nutrients in foods. The 1977-78 data were analyzed for food energy and 14 nutrients and other dietary components. For CSFII, 14 more components were added. Because of the considerable interest, three of these components--dietary fiber, tocopherol, and carotene--were added despite some limitations in the data.

One of the first steps in using, or in deciding whether to use, Continuing Survey data should be a review of the reports that have been published. Four reports are already out (CSFII Reports Nos. 85-1, 85-2, 85-3, and 86-1) and five others are in progress. These reports give an overview of the nature of the survey and present food and nutrient intakes by characteristics such as age, race, income, region, and urbanization. Food intakes are classified in 60 food groups and subgroups. Mean quantities of food eaten per individual per day are presented along with percentages of individuals who reported eating any food from the specified food group or subgroup.

The nutrient intake tables include data on mean nutrient intakes, mean intakes as a percentage of the 1980 RDA, mean intakes per 1,000 kilocalories, and the percentages of total food energy from protein, carbohydrate, and fat (total fat, saturated fat, monounsaturated fat, and polyunsaturated fat). Also presented are the frequency of eating; the nutrient contributions of snacks and of food obtained and eaten away from home; the percentages of individuals following special diets; and the percentages using vitamin and mineral supplements. Each report also contains an extensive set of tables describing the sample.

The next step is to obtain HNIS Administrative Report No. 378, which describes machine-readable data sets on composition of foods and results from food consumption surveys.

H.A. RIDDICK

Single copies of this report are available from the Human Nutrition Information Service (HNIS), 6505 Belcrest Road, Room 304, Hyattsville, Maryland 20782. The report gives ordering information, including prices, for the data sets (and for supporting materials on microfiche or paper) that are available from the National Technical Information Service (NTIS) of the Department of Commerce.

For each survey report, there is a tape that contains the survey results (including many variables not presented in the report) as well as tapes containing the nutrient data base used to calculate nutrient intake values. These tapes are organized into six files.

File 1 contains descriptive information about methodology and sample design that is basically taken from the report.

File 2 contains dataset characteristics such as the logical record length, blocksize, and the number of records. The file structure and format are also presented. CSFII data are in a hierarchical file containing four types of records for 1-day data tapes. Record type 10 contains basic characteristics of the household. Record type 20 contains data about each individual in the household. Record type 30 provides considerable detail about foods reported by all individuals who were age-eligible and participated in the survey. One record type 30 appears for each food item the individual reported. In addition to the amount consumed (in grams) and the nutritive values supplied by that amount of food, each food item contains descriptors such as time of day, name of eating occasion, a 7-digit USDA food code, whether salt was added at the table, and whether the food was obtained and eaten away from home. Record type 40 contains the sum of the nutrient values from all of the foods on record type 30 for each individual.

The data file format describes each variable on the tape by record type and column location, indicating whether there are implied decimals for the data field. The variable description is particularly useful since it repeats the entire question used to obtain the data, designates the question number and schedule used (screener, household, or individual), and indicates acceptable values, along with value labels, for each variable. Control counts for selected variables are contained in File 3. For categorical variables, such as race, the counts are the number of responses in each category. For continuous variables, such as last year's income, the counts are the number of responses within a given range. Also given are the counts of "don't know" and "no answer" responses. The file is intended to serve as a point of reference for those doing their own analysis of the data. The counts are also useful in helping to decide which variables (and variable categories) to include in an analysis plan.

File 4 contains the actual data. I will skip File 5 until I have described File 6.

File 6 is the Manual of Food Codes for Individual Intake. It is composed of approximately 5,000 7-digit food codes used for coding individual intake in CSFII. For each code, the manual includes a description of the food or beverage, common measures or weights of edible portions, conversion of measures and weights to grams, and a portion size used when the amount consumed was not specified.

Accompanying the manual on the CSFII tape are a list of food groups as categorized by the first three digits of the food code to serve as an aid to location of food groups and individual food codes; a list of notes and abbreviations used in the manual; and the coding guidelines used for CSFII. Specific guidelines include those for coding salads, fast-food sandwiches, and fast-food breakfast sandwiches.

File 5 contains the 7-digit food codes found in File 6 with an abbreviated (51 characters or less) version of the description for each code. This file does not contain all of the information found in the descriptions in the manual. It is provided as a time-saver for users. The 51-character description file can be linked to the data file when generating listings of individuals' food intakes or to other listings when exact identification of the food items is not necessary.

The microfiche that accompanies each data tape includes identical information with one

USING USDA CONTINUING SURVEY DATA

exception--File 4. File 4 of the tape includes the actual data, whereas Chapter 4 of the microfiche includes the interviewer instructions, food instruction booklet, screening forms, and questionnaires.

The interviewer instructions cover the following information: finding the specific geographic areas to be surveyed; identifying sample households; screening households; and obtaining cooperation from eligible households. There are general interviewing instructions as well as specific question-by-question instructions.

The Food Instruction Booklet (FIB) was developed to guide interviewers in collecting individual intake information. Anyone planning to use data from the CSFII should become familiar with it. In the interview, respondents are asked to give complete descriptions of the foods and beverages consumed the day before the interview and the actual quantities consumed. These descriptions must be complete and precise in order to make full use of the detailed set of food codes available. The FIB contains the specific questions or probes that should be asked for each food or food type. Copies of the microfiche may be ordered separately for those more interested in the printed support materials than in the survey data.

In addition to tapes of survey results, there are tapes of the nutrient data bases that correspond to the particular food consumption surveys. For example, Release 2 of the USDA Nutrient Data Base for Individual Food Intake Surveys was used to process the first day of intakes from the 1985 core monitoring group. This release 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 for the Southwestern United States. Release 2.1 covers the entire 1985 CSFII.

Three data sets are used to create each nutrient data base release. The Recipe File is a data set that controls the generation of the survey nutrient data base using the Primary Nutrient Data Set (PDS) and the table of retention factors. In the recipe file, each survey code is linked to one or more PDS items through a set of recipe codes. Links to single PDS items are treated as one-component recipes.

The amount of timely data available and the extensive documentation represent a major effort by USDA. A discussion of appropriate ways to interpret this data is beyond the scope of this talk. As a useful first step, I would suggest a review of a recent report by the Life Sciences Research Office of the Federation of American Societies for Experimental Biology (FASEB). The title of the report is "Guidelines for Use of Dietary Intake Data". Copies are available at a cost of \$14.00 prepaid from the FASEB Special Publications Office, 9650 Rockville Pike, Bethesda, Maryland 20814.