

The USDA Food Frequency Study
Frances A. Larkin, Ph.D.
School of Public Health, University of Michigan

The Food Frequency Study was designed to develop a questionnaire that will describe an individual's usual diet for one year. The study is based on several assumptions: 1) that an individual is reasonably consistent in his or her choice of food; 2) that he or she is sufficiently aware of the foods eaten to be able to recall them; 3) that he or she can generalize about eating habits over time and in terms of groups of foods, not just single foods; and 4) lastly, that he or she is willing and able to read and write. The specific study questions are:

1. Are questions about food frequencies valid indicators of last year's diet? Validity will be measured by the degree of correspondence between food frequencies for the past year and 16 days of recall/records sampled from the same period.
2. What characteristics of respondents, and of dietary patterns, are associated with validity or with various aspects of lack of validity such as bias?

The study design requires each respondent to provide information five times over the course of one year, four times providing one-day recalls and three-day records and the fifth time providing food frequency data for the previous year. Because interviewing for each round of data collection takes three months, the field work will take about 15 months. The 24 hour recall in the first round was used primarily as a respondent training tool for the three-day records. It was retained in all rounds as a training reinforcer and to provide another day of data. Each respondent is also given measuring spoons and cups, a ruler and a set of eighth-inch thickness measures adapted from the NHANES models to help in estimating amounts.

We have completed collecting the recall/records and are in the fifth round - that is, we are now collecting the food frequency data. This is a progress report using data from the pretest of the food frequency instrument administered to all respondents prior to the first 24 hour recall/three-day record, and from two rounds of food records.

Methods

Sample

Of the 273 enrolled into the project, we hope to retain at least 200 respondents, aged 23-50, roughly equally divided among males and females, black and white. We attempted to get an equal representation of high school education and greater than high school education within each category but we could not achieve this in our age group and communities. Refusal rates ran around 70 percent in Ypsilanti and 50 percent in Ann Arbor, perhaps because we required participation for a year and because our initial interview took about two hours. Respondents were enrolled by

interviewers visiting the homes in census tracts selected because they were at least 25 percent black in the 1980 census. The first interview consisted of the food frequency interview, a 24 hour recall and a brief health and demographic questionnaire. The 3 day record was left to be picked up 3 or 4 days later. In the second, third and fourth rounds, the interviewers visited the respondents for the 24 hour recall and left the 3 day record form. They again returned after several days to pick it up, and review the record before leaving the respondent's home.

Questionnaire Development

The food frequency tool is composed of 113 foods or food groups, each on a separate slip of paper which the respondent sorts into frequency categories.

The initial work in defining food groups was completed for an earlier contract, comparing the intake of food groups for NFCS and NHANESI. The identification of these groups was based on the food group order already built into the data bases, nutrient composition and similarity of use among foods. In the next stage of the development of food groups, we used the 24 hour and 2 day record of NFCS, selecting a 30 percent subsample of the adults aged 23 to 74.

Basically, we analyzed the pattern of occurrence of the most frequently consumed foods, comparing two foods at a time, for individuals over three days. This identified whether the two foods were eaten together, eaten at different times, or eaten without affecting each other. To correct for the frequency of occurrence, standardized scores were computed. A high positive score indicated a stable complementary relation, such as bread and butter. A high negative score indicated a substitute relationship, such as skim milk and whole milk. Next in the development of food groups, we considered the nutritive value of the food groups. Substitute food groups with similar nutrient values were combined - for example, whole wheat and white bread.

Complementary foods were put into separate groups but used in the description of the complementary food. For example, CRACKERS is one group and CHEESE another, but CHEESE is described as "CHEESE, including on CRACKERS." Some foods are used only as complements, such as margarine or butter. Some foods are partially used as complements, such as milk used in coffee or on cereal. For butter and margarine, the complementary use was asked in a probe. Milk in coffee was asked in a probe, and as an independent food group for milk as a drink.

Some foods used in the same manner are very different nutritionally, regular and light beer are examples. Beer, unspecified, is sorted first, then separated into the regular and light by a probe.

Some foods are used very differently depending upon whether they are eaten singly or in mixtures. For example, tomatoes are eaten raw, cooked and in sauces. Prevalence information determined the foods for which this occurred most often. These foods were then presented to the respondent as

separate foods.

Typical serving size was asked for each food except lettuce and an OTHER VEGETABLES group. A group of vegetables of different size and shape is difficult to generalize about in regard to serving size.

The frequency of use of food groups, whether the groups are used alone or in combination, the nutrient value, and the ease in generalizing about serving size were all considered in defining 110 food groups for the pretest in July through September, 1984. After three rounds of food records were collected, we revised the list, adding avocado, soy products and yogurt as separate foods and adding croissants to the bread group.

Food Frequency Questionnaire Administration

The basic format of the food frequency is a set of 113 slips of paper, 5 1/2 by 8 1/2, listing the name of the food or food group on the front and a partly precoded recording form on the back.

The interviewer places on the table before the respondent three envelopes marked with the frequencies NOT LAST YEAR, LESS THAN 12 TIMES, LAST YEAR, and 12 TIMES OR MORE LAST YEAR and says, "Please take these slips with the food printed on them and sort them into one or the other of three piles according to how often you ate the food last year." The NOT LAST YEAR and LESS THAN 12 TIMES piles are then placed in their respective envelopes and set aside. The slips sorted into the LESS THAN 12 TIMES pile will be used analytically in a diversity index but nutrients contributed by these rarely eaten foods will not be included in the determination of the whole year's intake.

The interviewer next takes the slips in the 12 TIMES OR MORE pile and hands them to the respondent saying "Now will you sort these foods which you said you ate 12 or more times last year into one of two piles according to whether you ate the food more often in CERTAIN SEASONS or about the SAME ALL YEAR ROUND." She then places the two envelopes so marked on the table. When this sort is completed, the slips in the CERTAIN SEASONS pile are placed in their envelope and temporarily set aside.

The interviewer then hands the SAME ALL YEAR ROUND slips to the respondent and says, "Now will you sort these foods which you said you ate about the same all year round into one of these piles according to how often you ate them last year." She then places envelopes marked with these frequencies on the table: " MORE THAN ONCE A DAY, ONCE A DAY, ABOUT 5-6 TIMES A WEEK, ABOUT 3-4 TIMES A WEEK, ABOUT 1-2 TIMES A WEEK, ABOUT 1-3 TIMES A MONTH. When the respondent finishes this sort, the interviewer puts the slips in the appropriate frequency envelope. Then, starting with the most frequent envelope, the interviewer asks, for each slip in turn, the amount of usual serving size for that food and any further probes indicated on the back of the slip. Figure 1 is the basic format. For example, the interviewer says, " You said you ate eggs in salads and sandwiches 3-4 times a week all year round. When you eat eggs in salads

and sandwiches, about how many do you eat?" All measuring aids are on the table to help in the estimating. The interviewer records directly on the back of the slip the frequencies, amounts and all other information and files the completed slip in a COMPLETED envelope.

When all the SAME ALL YEAR ROUND foods have been probed for amount and other information, the interviewer asks the respondent to sort the MOSTLY IN CERTAIN SEASON foods into the same frequency categories but for the appropriate season. She will then say, "You said you ate strawberries mostly in certain seasons and about 5-6 times a week in that season. About how much did you eat each time you had strawberries? What is your strawberry season? Did you eat strawberries at all during the rest of the year?" If yes, "How often?" The question about the length of the season was added after the pretest disclosed variable seasons for locally produced fruits and vegetables and for chili, cake and other items not ordinarily considered "seasonal." The question about seasonal foods being eaten during the rest of the year was included to get a closer estimate of annual intake.

Other Data Collected

At the first contact, interviewers collected demographic and health information, including information on the use of special diets for each of the health conditions described to the respondent. The demographic data includes employment status of the respondent and spouse and income. At the second through fifth contacts, a short life events questionnaire is completed. The respondent is asked about any changes in health, family composition, and other events that might affect eating patterns.

Results

Table 1 contains a preliminary analysis for only 172 participants, comparing the nutrient values from the pretest food frequency administered at the first round with eight recall/records collected from the first two rounds. The food frequency questionnaire was administered last July through September, asking about the previous year. The correct comparison, which is the heart of the study, will be between the food frequency currently being administered and the past year's 16 recall/records. We also look forward to seeing whether last year's food frequency bears any resemblance to the one being collected this year. That is, for those who say their diets did not change, what is the reliability of the instrument?

Total mean energy intake from the 24 hour recall, day 1, day 2 and day 3 records bear a close resemblance to each other, but are significantly different from results obtained by the food frequency. The frequency value is 20-25 percent greater than the means of the 24 hour and the records.

Is there any food group that is overestimated by the food frequency instrument? Since the 24 hour recall and records appear, at this stage of the analysis, to be similar, we combined the four days in order to compare information from the records and the food frequency.

We classified the foods from the food frequency and records into eleven food groups. For each method, the food frequency and the records, we computed the percentage of the total calories that each of these eleven food groups contributed.

Comparing the baseline food frequency and six months of record data for each food group, there is a moderate correspondence between percentages of calories from each food groups, especially for sugar by probe, beverages, fruits and dairy products, as shown here by correlation coefficients in Table 2. If these results hold up in the final analysis, it will indicate that most respondents were able to generalize fairly effectively about their intake of some foods. The paired t-test indicate that there is a significant difference between the two instruments for most food groups, excepting vegetables and fats.

The percentages of total calories contributed by grains, meats, desserts, sugar and beverages were systematically higher from the food frequency than the recall records. The percentages of total calories contributed by dairy products, eggs, fruits, and fats were systematically lower from the food frequency than the recall records.

In an attempt to understand the overestimation of meat and the underestimation of fruit, we have added a few questions to the food frequency currently being used. We asked about how many times a week meat as a group and fruit as a group are usually eaten, after the food frequency is completed. The problem of over and under estimation may be partially due to the number of foods or slips of paper in each food group. For example, there are 21 different meat items in the food frequency. Perhaps there are too many, so that small errors in each item add up to a substantial error. However, 24 different vegetable items are asked in the food frequency yet results from this instrument are quite similar to the recall/records. The egg group is the smallest, just two slips. The difference in results here may be because eggs are used in such a variety of foods that respondents may have trouble estimating them in a food frequency type format.

The length of time respondents took to complete the food frequency questionnaire is similar across race, age, and sex categories. The range of minutes in Table 3 is fairly large for some groups. Based on our observation of some respondents, the concept of how often food is eaten, and how much, is new to some people. The first part of the sort, whether or not food is eaten and whether or not it is a seasonal pattern, moves along fairly rapidly. Decisions on how often, and especially on how much, took a longer period of time. We plan to get an estimate of the distribution of time required for the different portions of the food frequency tool. We also plan to analyze whether questioning for the usual serving size is useful, or whether one can estimate intakes just as well using sex/age standards.

During the initial interview, about 30 percent of the respondents indicated that they were on a "special" diet, as listed in Table 4. About 25 percent of this group said they were on a weight loss diet. We do not

know the degree to which they altered their usual diet, or how long they followed the "special" diet.

Table 5 includes the foods that the respondents said were most affected by these special diets included meat, dairy products, sweets and vegetables. Meat was mentioned more often in this circumstance. The meat group was one of the groups that showed the greatest disparity between the food frequency and record recall measures of intake. There may be an ongoing process in the pattern of meat consumption.

When the three day record is picked up by the interviewer the record is reviewed and the respondent asked how typical these days have been. For various reasons, the kind or amount of food, how often, or the time the food was eaten, the records were not typical for a number of respondents. Table 6 lists the percentage of respondents reporting a typical diet for rounds one and two.

The use of a special diet, change in the use of a specific food and typicality of the diet may or may not affect the degree to which the food frequency and the recall record agree. Our analysis will explore these relationships.

Conclusion

A final caution against relying on any of the figures in the latter tables. Preliminary data are presented to illustrate trends in the analysis and the methods. There is also the more basic question of validity. We really don't know how valid the series of recall/records are. It could be that a series of recall/records is a good indicator of an individual's average diet, and that a food frequency is a good estimate of an individual's usual diet. We are assuming that the terms are interchangeable. By this time next year we should be able to give you some indication whether or not that assumption is valid.

Figure 1. Basic Format Food Frequency Questionnaire

1 Not last year
 2 12 times or more last year
 3 Less than 12 times last year

1 More seasonally
 2 Year round

Seasonal Frequency
 1 1+/day 2 1/day 3 5-6/wk
 4 3-4/wk 5 1-2/wk 6 1-3/mo

How much did you usually have? _____

 What is your _____ season? _____

(12) _____
 (AMT-13-17) _____
 (MEAS 18-19) _____
 (WKS 20-21) _____

Year round Frequency
 1 1+/day 2 1/day 3 5-6/wk
 4 3-4/wk 5 1-2/wk 6 1-3/mo
 7 1-11/yr

How much did you usually have? _____ (AMT 23-27)
 _____ (MEAS 28-29)
 _____ (12) _____ (MSU 30-35)

Did you eat/drink _____ at other times last year? NO YES: How often? _____

(5-9)	0	0	6	1	1
(10)					
(11)					
(22)					
(AMT 23-27)					
(MEAS 28-29)					
(MSU 30-35)					

Table 1. Mean Nutrient Intake Estimated from Food Frequency, Recall and Records, Preliminary Results

	Food Frequency I (N=172)	Rounds I, II Records (N=234)				Mean
		24 Hr	Day 1	Day 2	Day 3	
Energy (kcal)	2808	2201	2111	2196	2223	2183
Protein (gm)	103	81	81	81	81	81
Fat (gm)	127	96	93	95	97	95
Carbohydrate (gm)	305	242	229	244	239	239

Food Frequency significantly different (p < 0.1) from records by paired t-test

Table 2. Comparison of Percentage of Total Energy Contributed by Eleven Food Groups for Food Frequency and Recall plus Records: Preliminary Results

Food Group	Percent of Total Calories			Paired t-test Significance
	Food Frequency I	Recall and Records I & II	r	
Grains	23.8	19.3	.37	.0000
Meats	21.0	15.5	.40	.0000
Desserts	7.3	4.5	.44	.0000
Sugar (probe)	1.7	1.1	.69	.0001
Dairy	10.8	14.1	.56	.0000
Beverages	9.0	7.0	.66	.0000
Eggs	2.1	2.4	.42	.0402
Fruits	6.9	10.7	.61	.0000
Vegetables	7.6	8.2	.35	.1535
Fats (butter, margarine, probe)	4.5	4.3	.31	.6571
Other fats	6.3	7.5	.25	.0116

Correlations are all significant at $p < .01$

Table 3 Time (in Minutes) to Complete Baseline Food Frequency by Race, Sex and Age, Continuing Participants As of Round Four Only (N=234): Preliminary Results

Demographic Group	N	Time (in minutes)		
		Minimum	Maximum	Mean
White				
Men				
23-34	36	34	160	75
35-50	29	40	120	72
	65			
Women				
23-34	37	48	120	73
35-50	36	15	150	73
	73			
Black				
Men				
23-34	21	50	145	78
35-50	24	45	119	71
	45			
Women				
23-34	25	47	120	78
35-50	24	30	110	75
	49			
Missing information for time	2			
Total	234			

Table 4. Frequency of Special Diets Mentioned from Health Information Questions:

<u>Special Diets</u>	<u>Numbers of Times Mentioned</u>
Weight loss	17
Pregnant/lactating	8
Low salt	7
Low cholesterol/fat	5
Increased calories	4
Vegetarian	4
Lactose avoidance	4
Low salt and low fat	3
Diabetic	3
Hypoglycemia	2
Gallbladder	2
Migraine	2
High fiber	1
Low fat and high fiber	1
Other	7
Total Special Diets	70
Respondents	234

Table 5. Number of Continuing Participants as of Round Four Reporting Foods Eaten Less, More, or Not at All for Special Diets

Meat	27
Dairy Products	25
Sweets/Sugar	20
Vegetables	19
Fruits/Juices	15
Non-Sweet Grain Products	12
Fish	11

Twelve other food groups were mentioned by fewer than ten respondents.

Table 6. Percent Reporting Three-Day Record was Typical for Rounds One and Two, Continuing Participants as of Round Four Only (N=234)

Typicality of 3-Day Record for	Percent Reporting Typical	
	Round 1	Round 2
Kinds of foods	75	75
Amounts of foods	78	78
Frequency of eating	83	86
When food eaten	82	88