R84-62

Validity of telephoned diet recalls and records for assessment of individual food intake[1–4]

Nora J Krantzler, PhD, MPH, Barbara J Mullen, BS, Howard G Schutz, PhD, Louis E Grivetti, PhD, Cathy A Holden, MS, and Herbert L Meiselman, PhD

ABSTRACT Six methods of assessing individual food intake reported by telephone were compared for accuracy. A sample of 107 students eating in a dormitory dining hall was used. Two 7-day and four 3-day diet records were reported by telephone to either an interviewer or an answering device; and fourteen 6-h recalls and seven 24-h recalls were obtained by an interviewer over the telephone. To examine validity, the investigators observed respondents' intake for 28 days. Food item agreement scores comparing observed and reported data were calculated. Seven-day records were most accurate (87% food item agreement); 3-day records and 6-h recalls were equivalent (75%); and 24-h recalls were least accurate (69%). There was no significant difference in accuracy of reporting records to an interviewer versus a recording devvice, but respondents preferred the interviewer. Results of telephoned reporting compare favorably with personal interview techniques used by other investigators, indicating that telephoned methods should be further explored in community settings. *Am J Clin Nutr* 1982;36:1234–1242.

KEY WORDS Validity, telephone, diet recall, diet record

Introduction

Diet is implicated in the etiology of many diseases. Current epidemiological research has focused on relationships between diet and cardiovascular disease (1–3), and cancer of the breast (4), colon (4), pancreas (5), and stomach (6), in addition to problems of over- and undernourishment. Government agencies also maintain dietary surveillance on a national level. Accurate estimation of current dietary intake is crucial for these investigations when they have a prospective or cross-sectional research design, monitoring intake over a defined period of time. Recent research (7) suggests the potential usefulness of methods focusing on foods and food consumption patterns, as well as traditional methods of examining nutrient intake.

Nevertheless, a major problem with dietary methods remains: those assumed to be most accurate are least feasible and most reactive (producing changes in the diet). While having foods weighed and measured by a nutritionist may be the most accurate means of determining nutrient intake, it is usually reserved for use on metabolic wards or other small, select samples due to high cost and excessive burden on respondents (8). Further, because of reactivity, weighed intake may not be a valid way to study food *habits*. Therefore, most

[1]From the Departments of Nutrition and Consumer Sciences (NJK, BJM, HGS, LEG, CAH), University of California, Davis, CA 95616 and the Behavioral Sciences Division, Science and Advanced Technology Laboratory (HLM), Natick Research and Development Laboratories, US Army, Natick, MA 01760.

[2]Supported by the Food and Drug Administration Agreement W13G07-79246-009, 1980.

[3]Address reprint requests to: Herbert L Meiselman, Behavioral Sciences Division, Science and Advanced Technology Laboratory, Natick Research and Development Laboratories, US Army, Natick, MA 01760.

[4] Results from the food frequency questionnaires will be reported elsewhere.

Received February 22, 1982.
Accepted for publication May 21, 1982.

studies of "free-living" populations involve self-reporting by individuals, using methods that are easier to administer but less precise. Two of these are the diet recall, an interview in which respondents are asked to remember foods consumed, and the diet record, requiring respondents to keep written accounts of their consumption.

Variations on diet recall and diet record are among the most frequently used self-report techniques for assessing food intake of individuals or groups over specific time periods. Problems with the recall technique, specifically the 24-h recall, have been noted due to its usual method of single administration (9), validity (10), and reliability (11). Diet records are more burdensome for respondents and thus have more potential for reactivity. In addition, as traditionally administered, both of these methods require respondents to visit with a professional interviewers in the respondents' homes or in a clinic setting; this may be burdensome, intrusive, or both, resulting in high expense and administrative difficulties when using such methods on large populations. Because of these issues, there is currently much interest in the potential use of the telephone for obtaining dietary data. A recent study (12) investigated the possibility of obtaining diet records by telephone, to facilitate data collection and reduce expenses, and demonstrated that valid data on nutrient intake could be obtained in this manner.

The study reported here was designed to test the validity of several different dietary methods that use telephone reporting. Herein we evaluate food intake methodologies administered by telephone: two diet recall and four diet record variations. Records collected using a voice-activated telephone recording device are compared with those obtained by an interviewer. Our concern in the present analysis is to determine accuracy of reporting of food items consumed, for potential use in analyzing food consumption patterns of individuals

Methods

Sample

A convenience, volunteer sample of 107 students from a single dormitory complex at the University of California, Davis, agreed to participate. Eighty-six percent were lower classmen (freshmen and sophomores); all were between the ages of 17 and 25 yr. Sixty-four percent were female, and 36% were male. Distribution by academic major was as follows: life sciences 24.1%, social sciences 18.2%, physical sciences 11.7%, arts and humanities 6.6%, food or nutrition-related majors 4.4%, other combinations of majors 4.4%, and undeclared 30.7%. The predominant self-reported ethnic background was Caucasian-European; only 3% of participants were foreign students. Eight percent reported that they were vegetarians. Participants were randomly assigned to method group in accord with the design in **Figure 1.**

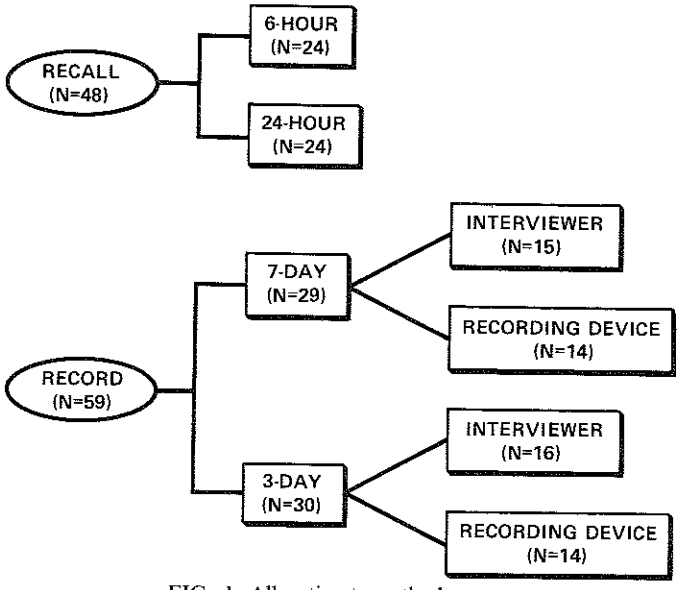

FIG. 1. Allocation to method groups.

Validity criterion

We required a validity criterion that was accurate, easy to administer, inexpensive, and unobtrusive (nonreactive). In the experimental setting of the dormitory dining hall, daily menus designed for a 6-wk cycle were made available to us in advance. We used these menus to develop an indirect observation technique that was easy to administer and unobtrusive. Participants were given a 5 × 7-inch printed form at the beginning of each meal listing all foods available at the meal.

During the meal, they marked food selected and number of servings taken. To lessen the reactivity of the method, respondents were asked to rate each food item selected using a nine-point preference scale provided to the right of each food listed. In this way, collection of consumption data was deemphasized. After the meal, the forms were attached to the respondents' tray and returned to the kitchen area on a conveyor belt. Project staff members removed all tray with forms, checking each form for accuracy using visual estimation.[5] Data on 28 breakfasts, 19 lunches, and 26 dinners were available and collected during a 28-day period, May to June 1980. Only foods monitored during these meals were considered in the analysis; we were not able to monitor foods consumed away from the dining commons.

Diet recalls

Six- and 24-h recalls were obtained from two groups of participants. All interviews were conducted between 6:30 and 9:30 PM. In one group, fourteen 6-h recalls were administered by telephone during the 28-day observation period. Distribution included 10 weekdays and 4 weekend days. Participants were telephoned by an interviewer and asked to recall all foods and beverages consumed in four different 6-h periods within the preceding 24 h (12:00 midnight to 6:00 AM, 6:00 AM to 12:00 noon, 12:00 noon to 6:00 PM, and 6:00 PM to 12:00 midnight), alternated during the 14 days surveyed. In the second group, seven 24-h recalls were administered throughout the month, representing each day of the week. This group of participants was asked to recall all items consumed during the 24-h period from 12:01 AM to midnight the preceding day. For both groups, then, there was a minimum time lapse of 18.5 h and a maximum of 45.5 h between consumption and reporting.

Diet records

Three- and 7-day records were obtained from two groups of participants. Each received detailed instructions and a sample diet record in the mail. Forms for writing records on were included in the packet. Four 3-day records were obtained, totalling 12 days. Each record-keeping cycle began on a different day, with cycles ranging from 2 to 6 days apart. Record-keeping began at 12:01 AM on day 1, and ended at midnight on day 7. Each day, respondents telephoned their record for the previous day to an interviewer (n = 16), or dictated records to a telephone recording device (n = 14). The next day's record began where the previous day's left off. Seven-day records, spaced 1 wk apart, were kept two times, for a total of 14 days. As with 3-day records, participants telephoned their records to either an interviewer (n = 15) or a recording device (n = 14).

Statistical methods and data analysis

Values are expressed as mean ±SD and as percentages. Differences between percentages are tested for statistical significance using the difference of proportions test (13). To test relationships between variables, Pearson's product moment coefficients are calculated (13).

Validity is assessed using a food item agreement score, expressed as number of foods correctly identified/total number of foods reported × 100. Thus, we compare the methods for validity by examining food item agreement of observational and methods data, by meal (breakfast, lunch, and dinner), by food group, and by individual. We also evaluate errors participants made in estimating their intake—specifically, over reporting (foods reported but not observed) and under reporting (foods observed but not reported).

Results

Participation

Participation varied among the method groups to which respondents were assigned. The highest withdrawal rates occurred in the diet record groups. 17% (n = 5) of the 7-day record sample, and 20% (n = 6) of the 3-day record group, withdrew from the study. In contrast, none withdrew from the 6-h recall group, and 8% (n = 2) dropped from the 24-h recall group. A total of 55 and 58% of the expected 6- and 24-h recalls, respectively, were completed, compared with 44 and 27% of the expected 3- and 7-day records.

Recalls

Food item agreement by meal for 6- and 24-h recalls is presented in **Table 1**. For each meal, the 6-h recall yields a higher validity score than the 24-h recall. When all meals are pooled, food item agreement scores for 6- and 24-h recalls are 75 and 69%, respectively (**Figure 2**). These differences in food item agreement are not statistically significant.

Table 2 shows validity scores for each survey method for 14 different food groups. Validity scores for 6-h recalls range from 53.3% (nuts/seeds) to 100% (soups). For 24-h recalls, scores range from 33.3% (soups) to 87.5% (sugar/sweeteners)—lower than 6-h recall scores on both ends of the scale. Scores on condiments and soups are significantly higher for the 6-h recall group. For sand-

[5] The accuracy of visual estimation of plate waste has been demonstrated in a recent study by Comstock et al., J Am Diet Assoc 1981;79:290–6.

TABLE 1
Validity of 6- and 24-h recalls by meal

	Breakfast				Lunch				Dinner			
	6-h (n = 20)		24-h (n = 18)		6-h (n = 22)		24-h (n = 21)		6-h (n = 19)		24-h (n = 23)	
	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*
No foods per meal	4.74	1.95	5.33	2.19	5.00	2.29	5.22	1.99	6.67	2.51	6.77	2.48
No foods agreeing	3.95	1.93	4.04	2.06	3.66	2.02	3.37	1.89	4.67	2.58	4.40	2.23
Foods observed, not reported	0.79	1.32	1.21	1.13	1.34	1.39	1.83	1.46	2.00	1.82	2.36	2.10
Food reported, not observed	0.55	0.69	0.83	1.09	0.63	0.86	1.32	1.37	1.33	1.28	0.97	1.07
Total food entries*	180		277		205		214		280		494	
Total meal entries	38		52		41		41		42		73	
Percent agreement	83%		76%		73%		65%		70%		65%	

* Foods may be listed more than once.

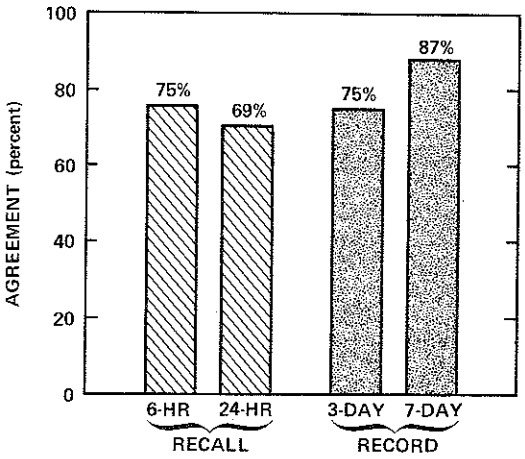

FIG. 2. Validity of recalls and records, all meals combined.

wiches, however, the 24-h recall produces a significantly higher score. Differences between other food groups are not significant. Overall, sugar/sweeteners (87.5 and 80%) and dairy foods (83.3 and 82.6%) were recalled correctly most often. Fats/oils, meat/fish, beverages, and casseroles were recalled correctly at least 75% of the time. Least well-recalled foods include nuts/seeds (61.5 and 53.3%), snacks/desserts (58.0 and 58.3%), and condiments (52.8 and 68.4%).

Food item agreement scores were calculated for each participant, using a combined score across all foods. Scores range from 27 to 100%, with 48% of individuals recalling correctly between 70 to 89% of all foods (**Figure 3**). When scores are averaged across individuals, the mean and SD for food item agreement are 77 ± 16% for 6-h recalls, and 70 ± 14% for 24-h recalls.

Records

Food item agreement by meal for 3- and 7-day records is presented in **Table 3**. For each meal, scores computed from 7-day records are higher than those from 3-day records. For all meals combined, there is a mean agreement score of 87% for 7-day records, compared with 75% for 3-day records. These differences are not statistically significant.

Analysis by food groups (Table 2) shows that 3-day scores range from 49.3% (condiments) to 100% (sandwiches, soups). The scores for 7-day records show a low of 67.3% (condiments) and a high of 100% (casseroles). The 7-day record produces significantly higher values for grains/starches, berries/fruits, condiments, casseroles, and soups. For the two record types combined, sandwiches, meat/fish, and dairy foods are reported correctly more than 90% of the time. Least accurately reported are condiments (49.3% and 67.3% for 3- and 7-day records, respectively).

Food item agreement scores for individual participants for all foods combined range from 56 to 100% for 3-day records; for 7-day records, scores range from 61 to 100%. The distribution of scores is presented in Figure 3. When averaged across individuals, the mean and SD for the agreement scores are 81 ±15% for 3-day records and 86 ±8% for 7-day records.

Table 4 presents food items agreement for

TABLE 2
Validity by food groups—all methods; percentages of food items accurately reported

Food group	6-h recall	vs	24-h recall	3-day record	vs	7-day record
Meat/fish	85.5% (55)*		70.4% (54)	94.4% (72)		96.0% (201)
Dairy	82.6% (109)		83.3% (174)	89.1% (192)		91.0% (335)
Fats/oils	76.9% (26)		81.1% (37)	79.5% (39)		89.9% (68)
Grains/starches	77.0% (87)		64.7% (133)	77.2% (189)		90.7% (377)†
Vegetables/legumes	69.0% (71)		64.4% (104)	80.9% (89)		89.5% (152)
Nuts/seeds	53.3% (15)		61.5% (13)	77.8% (9)		76.9% (13)
Berries/fruits	63.2% (57)		59.8% (102)	63.1% (84)		86.0% (165)†
Beverages	81.5% (92)		72.4% (105)	84.4% (147)		87.1% (203)
Snacks/desserts	58.3% (48)		58.0% (50)	72.9% (59)		82.2% (118)
Condiments	68.4% (57)		52.8% (106)‡	49.3% (73)		67.3% (179)‡
Sugar/sweeteners	80.0% (5)		87.5% (8)	83.3% (12)		90.5% (21)
Sandwiches	58.8% (18)		73.1% (26)	100.0% (20)		94.1% (17)
Casseroles	70.0% (10)		78.6% (14)	72.7% (11)		100.0% (17)‡
Soups	100.0% (7)		33.3% (9)	100.0% (11)		71.4% (28)

* n given in parentheses = no foods per group. Percentages reflect all food items reported in each food group.
† $p < 0.01$.
‡ $p < 0.05$.

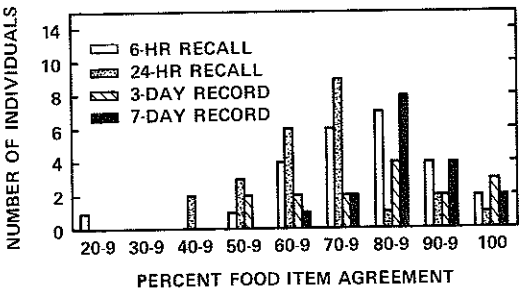

FIG. 3. Distribution of food item validity scores.

dietary records telephoned to an interviewer, compared with those collected using an answering device. Overall, food item agreement for records telephoned to an interviewer is 85% compared with 84% for records called to an answering device. For 3-day records, 79% of food items reported to an interviewer and 82% of those recorded on the answering device are in agreement. Food item validity for the 7-day record shows 89 and 85% agreement for foods reported to interviewer and recording device, respectively. These percentages are not significantly different.

To examine whether or not validity decreased as more daily records were kept, we compared food item agreement scores with the total number of records or recalls obtained for each participant. We found a significant correlation for the 3-day records only; ie, as the number of records increases, the percent of food item agreement decreases ($p < 0.05$).

We also tested the hypothesis that individuals on a special diet might remember or record their intake more accurately than those who were not dieting. The respondents in the recall groups who claimed to be dieting ($n = 6$) have a mean agreement score of 81%, while the other respondents have a mean score of 77%. These differences, however, are not statistically significant.

Errors in reporting

Across methods, foods tend to be underreported (observed but not reported) rather than overreported (reported but not observed). When summarized by method, underreporting of foods ranges from 13% for 7-day records to 31% for 24-h recalls. Overreporting of foods ranges from 10% for 7-day records to 18% for 24-h recalls (Data not shown). However, these differences are not statistically significant.

Examination of error by food group reveals large differences between food groups (**Table 5**). Overall, recalls produce a significantly higher percentage of errors than records for four groups: meat/fish/poultry, vegetables/legumes, berries/fruits, and beverages. Underreporting errors are more frequent for recalls than for records, with the exception of sugar/sweeteners. Comparing 6- and 24-h recalls for underreporting shows that 24-h re-

TABLE 3
Validity of 3- and 7-day records by meal

	Breakfast				Lunch				Dinner			
	3-day (n = 13)		7-day (n = 18)		3-day (n = 12)		7-day (n = 15)		3-day (n = 14)		7-day (n = 19)	
	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*	$\bar{x}$	*SD*
No foods per meal	5.43	1.87	5.14	2.08	5.55	2.00	5.72	1.97	6.81	6.31	6.34	2.24
No foods agreeing	4.33	1.96	4.66	1.99	4.12	1.90	5.00	1.90	4.90	2.38	5.31	2.04
Foods observed, not reported	1.08	1.36	0.48	0.84	1.24	1.45	0.72	0.94	1.19	1.48	0.99	1.24
Foods reported, not observed	0.66	0.94	0.47	0.81	0.90	1.21	0.63	1.09	1.19	1.49	0.57	0.89
Total food entries*	434		709		233		475		429		850	
Total meal entries	80		138		42		83		63		134	
Percent agreement	80%		91%		74%		87%		72%		84%	

* Foods may be listed more than once.

TABLE 4
Record validity, answering device vs interviewer

Method	Answering device (n = 14)			Interviewer (n = 20)		
	Agreement*	$\bar{x}$ error†	Meals no	% Agreement	$\bar{x}$ error†	Meals no
3-day record	82	1.02	51	79	1.03	135
7-day record	85	0.66	180	89	0.62	175
3- and 7-day records combined	84	0.74	231	85	0.79	309

* Percentage agreement = (no foods correctly reported/no foods observed) × 100.
† (no foods observed but not reported) + (no foods reported but not observed).

TABLE 5
Errors in reporting: recalls vs records (summary)

	% underreported*					% overreported†				
	Recalls	n	Records	n	p‡	Recalls	n	Records	n	p
Meat/fish/poultry	22.0	109	4.4	273	<0.001	15.6	109	3.7	273	<0.001
Dairy	16.8	273	9.7	527	0.01	13.6	273	12.0	527	NS
Fats/oils	20.6	63	19.6	107	NS	36.5	63	23.4	107	NS
Grains/starches	30.5	220	13.8	566	<0.001	17.3	220	12.0	566	NS
Vegetables/legumes	33.7	175	13.7	241	<0.001	23.4	175	13.7	241	0.05
Nuts/seeds	42.9	28	22.7	22	NS	14.3	28	22.7	22	NS
Berries/fruits, etc	38.4	159	21.7	249	<0.001	25.8	159	10.0	249	<0.001
Beverages	23.4	197	14.0	350	0.01	14.7	197	6.6	350	0.01
Snacks/desserts	41.8	98	20.9	177	<0.001	24.5	98	16.9	177	NS
Condiments	41.7	165	38.1	252	NS	14.5	165	15.1	252	NS
Sugar/sweeteners	15.4	13	61.5	33	0.01	61.5	13	60.6	33	NS
Sandwiches	32.6	43	2.7	37	0.01	7.0	43	2.7	37	NS
Casserole type	25.0	24	10.7	28	NS	0.0	24	3.6	28	NS
Soups	37.5	16	20.5	39	NS	18.8	16	10.3	39	NS

* No foods observed but not reported/total no foods observed in that food class × 100.
† No foods reported but not observed/total no foods observed in that food class × 100.
‡ Test for differences between proportions for uncorrelated data on small samples.

calls produce more frequent errors for 8/14 food groups present (**Table 6**). These differences are significant for three food groups: grains/starches, condiments, and soups. Six-hour recalls yield more underreporting errors than 24-h recalls for 6/14 food groups; none of these differences is statistically significant. Comparing 3-day and 7-day records indicates that 3-day records produce more frequent underreporting errors than 7-day records, ex-

TABLE 6
Errors in reporting

Food group	% underreported*						% overreported†					
	6-h	24-h	p‡	3-day	7-day	p	6-h	24-h	p	3-day	7-day	p
Meat, fish, poultry	14.5	29.7	NS	5.6	4.0	NS	10.3	20.4	NS	2.8	4.0	NS
	(55)	(54)		(72)	(201)		(55)	(54)		(72)	(201)	
Dairy foods	17.4	16.7	NS	10.9	9.0	NS	16.5	10.3	NS	13.5	11.0	NS
	(109)	(174)		(192)	(335)		(109)	(174)		(192)	(335)	
Fats, oil	23.1	18.3	NS	20.5	19.1	NS	15.4	51.4	0.01	33.3	17.6	NS
	(26)	(37)		(39)	(68)		(26)	(37)		(39)	(68)	
Grains, starches	22.4	35.3	0.05	22.8	9.3	<0.001	19.5	15.8	NS	18.5	8.8	<0.001
	(87)	(133)		(189)	(377)		(87)	(133)		(189)	(377)	
Vegetables, legumes	30.4	35.6	NS	19.1	10.5	NS	21.1	25.0	0.05	12.4	14.5	NS
	(71)	(104)		(89)	(152)		(71)	(104)		(89)	(152)	
Nuts, seeds	46.7	38.5	NS	22.2	23.1	NS	13.3	15.4	NS	11.1	30.8	NS
	(15)	(13)		(9)	(13)		(15)	(13)		(9)	(13)	
Berries, fruits	36.8	40.2	NS	36.9	13.3	<0.001	17.5	30.4	NS	16.7	6.7	0.05
	(57)	(102)		(84)	(165)		(57)	(102)		(84)	(165)	
Beverages	18.5	27.6	NS	15.6	12.8	NS	3.8	19.0	0.01	9.5	4.4	NS
	(92)	(105)		(147)	(203)		(92)	(105)		(147)	(203)	
Snacks, desserts	41.7	42.0	NS	27.1	17.8	NS	12.5	36.0	0.01	25.4	12.7	0.05
	(48)	(50)		(59)	(118)		(48)	(50)		(59)	(118)	
Condiments	31.6	47.2	0.01	50.7	33.0	0.01	14.0	15.1	NS	27.4	4.5	<0.001
	(57)	(106)		(73)	(179)		(57)	(106)		(73)	(179)	
Sugar, sweeteners	20.0	12.5	NS	16.7	9.5	NS	80.0	50.0	NS	41.7	71.4	NS
	(5)	(8)		(12)	(21)		(5)	(8)		(12)	(21)	
Sandwiches	41.2	26.9	NS	0.0	5.9	NS	0.0	11.5	NS	0.0	5.9	NS
	(17)	(26)		(20)	(17)		(17)	(26)		(20)	(17)	
Casseroles	30.0	21.4	NS	27.3	0.0	0.05	0.0	0.0	NS	0.0	5.9	NS
	(10)	(14)		(11)	(17)		(10)	(14)		(11)	(17)	
Soups	0.0	66.7	0.05	0.0	28.6	NS	0.0	33.3	NS	27.3	3.6	NS
	(7)	(9)		(11)	(28)		(7)	(9)		(11)	(28)	

* No foods observed but not reported/total no foods observed for that food class × 100.
† No foods reported but not observed/total no foods observed for that food class × 100.
‡ Test for difference between proportions for uncorrelated data on small samples.

cept for 3 food groups (nuts/seeds, sandwiches, and soups) (Table 6). For these three groups, 7-day records have a greater percentage of errors, but none is statistically significant. The higher percentage of errors for 3-day than for 7-day records is significant for grains/starches, berries/fruits, condiments, and casseroles.

Errors of overreporting are more frequent for recalls than for records, for all but one food group (casseroles) (Table 5). Recalls produce a significantly greater percent of overreporting errors than records for four food groups: meat/fish/poultry, vegetables/legumes, berries/fruits, and beverages. Comparison of 6- and 24-h recalls for errors of overreporting shows that 24-h recalls produce more errors for 10/14 food groups (Table 6). These differences are significant for fats/oil, vegetables/ legumes, beverages, and snacks/ desserts. Six-hour recalls show a higher percentage of errors for three food groups, but none is significant. The two types of recalls are equivalent for one food group (casseroles), which are never overreported. Comparing 3- and 7-day records, we find that 3-day records have a higher percentage of overreported foods for 8/14 groups. These differences are significant for four: grains/starches, berries/fruits, snacks/desserts, and condiments. Overreported items in 7-day records are more frequent for 6/14 food groups, although none of these differences is statistically significant.

Discussion

Results of this study demonstrate that it is feasible to collect accurate data on food intake using the telephone. Our data suggest the 7-day record technique provides better validity results than the 3-day record, 6-h recall, or 24-h recall. Accuracy of data from 3-day records and 6-h recalls is approxi-

mately equivalent; 24-h recalls exhibit a lower but not significant validity score. Overall, we find that records are more accurate than recall.

While shorter, more frequent periods of data collection yield higher validity scores than longer, less frequent periods, they are not necessarily easier for respondents to complete. This is shown by the lower rate of completion by those remaining in the study (27% compared with 44%). We suggest two possible explanations. First, considering 3-day records, repeated starting and stopping of record-keeping may have been too burdensome for some respondents. Those in the 6-h recall group, once reached by telephone for the interview, may have preferred to give more information at a time, with fewer contacts. A second but perhaps related reason has been suggested by Sharp and Frankel (14), who recently found that the perceived importance and applicability of survey results is more critical than survey length of time in affecting response rates and attitudes. Perhaps respondents who were asked to provide information over longer periods of time believed they were making a more important contribution than those providing less information each time. This suggests that a method such as the 7-day record, normally considered quite burdensome, could be used by the general population if respondents are convinced of the usefulness of the research, and if they are given support and encouragement over the telephone. Shorter but more frequently administered methods, while more accurate in determining foods consumed, may be more burdensome than longer, less frequent assessments.

Use of the telephone answering device for reporting diet records, while producing data approximately as accurate as those telephoned to an interviewer, appears to be more difficult or burdensome for respondents than talking with the an interviewer. As with 3- and 7-day records, this is shown by lower completion rates. Good rapport may be established on the telephone by the interviewer, particularly when respondents are contacted repeatedly.

We believe that completion of the diet records was affected by the requirement that respondents initiate the telephone calls themselves, to report their intake. This was the method followed by Raker (12), who also compared interviewer and answering device techniques (the latter requires respondents to take the initiative in calling). In a subsequent community study conducted as part of the Stanford Heart Disease Prevention Program, interviewers telephoned respondents and obtained substantially improved completion rates (Krantzler NJ, et al, unpublished data).

Analysis by food group reveals that foods eaten regularly, or those contributing the major part of the meal, are better reported. Dairy and meat/fish groups show the highest overall reporting accuracy. The dairy category is comprised mainly of milk; as milk tends to be consumed frequently and at regular times by most individuals, perhaps this consistency enables respondents to better report their intake. On the other hand, foods such as condiments and nuts/seeds are less likely to be reported accurately. This may be because such foods are not usually central to a meal and are generally eaten in small amounts as garnishes to main courses or as snacks. Hence, they may be underreported more often than other foods.

Errors of overreporting show less variation among food groups than errors of underreporting, with the exception of sugar/sweeteners. This category is overreported four to five times more often than any other group. It may be that sugar is added to foods and beverages in a routine or unconscious way, so that respondents may have forgotten to mark it on the observation ("preference") form, but did report it when asked more rigorously about their intake. Such differences in reporting among different food groups suggest a need to remind respondents to report as completely as possible all foods eaten, and they illustrate the need for creating specific protocols to probe for least well-reported items if those are of particular interest.

Validity scores for each respondent for all foods combined show moderate variation within method groups in ability to accurately recall or record one's diet. When averaged across individuals, the mean and SD for the agreement scores are 77 ± 16% for 6-h recalls, 70 ± 14% for 24-h recalls, 81 ± 15% for 3-day records, and 86 ± 9% for 7-day records. These results compare favorably with those

reported by Schnakenberg et al (15), who found food item agreement scores of 80% using a combination diary-personal interview technique. The slight downward trend in validity by meal from breakfast to lunch and dinner may reflect the more routine and easily identified nature of most breakfast foods served in the dining hall. Since this study is based on a self-selected sample of students eating in a dormitory dining hall, results may not be generalizable to other population groups. The institutional environment provides a valuable setting for observing individuals' food habits in an unobtrusive, nonreactive way. However, the wider applicability of telephoned methods remains to be tested in community settings.[6]

The authors thank Cynthia Hoopes and Alexander Lamb for assistance with data processing, and Elizabeth M. Comstock for helpful critiques of data analysis and presentation.

References

1. Gordon TA, Garcia-Palmieri M, Kannel WB, et al. Diet and its relation to coronary heart disease and death in three populations. Circulation 1981;63:500–15.
2. Hulley SB, Sherwin R, Nestle M, Lee PR. Epidemiology as a guide to clinical decisions—II. Diet and coronary heart disease. West J Med 1981;135:25–33.
3. Shekelle RB, Shryock AM, Paul O, et al. Diet, serum cholesterol, and death from coronary heart disease. N Engl J Med 1981;304:65–70.
4. Correa P. Epidemiological correlations between diet and cancer frequency. Cancer Res 1981;41(part 2):2685–9.
5. MacMahon B, Yen S, Trichopoulos D, Warren K, Nardi G. Coffee and cancer of the pancreas. N Engl J Med 1981;304:630–3.
6. Hankin J, Rhoads G, Glober G. A dietary method for an epidemiologic study of gastrointestinal cancer. Am J Clin Nutr 1975;28:1055–61.
7. Schwerin HS, Stanton JL, Riley AM, et al. Food patterns and health: a reexamination fo the Ten-State and Hanes I surveys. Am J Clin Nutr 1981;34:568–80.
8. Burk M, Pao E. Methodology for large-scale surveys of household and individual diets. Washington, DC: USDA Home Economics Research Report No 40, 1976.
9. Garn S, Larkin F, Cole P. The problem with one day dietary intakes. Ecol Food Nutr 1976;5:245–7.
10. Madden P, Goodman SJ, Guthrie HA. Validity of the 24-hour recall. J Am Diet Assoc 1976;68:143–7.
11. Balogh, M, Kahn H, Medalie J. Random repeat 24-hour dietary recalls. Am J Clin Nutr 1971;24:304–10.
12. Raker MR. The validity of a telephoned food record. Master's thesis, Pennsylvania State University Graduate School, Department of Nutrition, 1979.
13. Snedecor GW, Cochran WG. Statistical methods. Ames, IA: The Iowa State University Press, 1972.
14. Bureau of Social Science Research. Long interviews are not main cause of refusals. BSSR Newslett 1980; Fall:1–2.
15. Schakenberg DD, Hill TM, Kretsch MJ. Diary-interview technique to assess food consumption patterns of individual military personnel. Assessing changing food consumption patterns. Washington, DC: Food and Nutrition Board, National Research Council, National Academy Press, 1981:180–97.

[6] Based on these results, three of these methods are currently being tested in four California communities, as part of the Stanford Heart Disease Prevention Program, with support from the Food and Drug Administration and the US Army, Natick.