Vital and Health Statistics

Dietary Methodology Workshop for the Third National Health and Nutrition Examination Survey

Series 4:
Documents and Committee Reports
No. 27

The Dietary Survey Methodology Workshop (March 16–18, 1986) was sponsored by the National Center for Health Statistics for the purposes of reviewing, evaluating, and making recommendations for the selection of dietary methodologies for the third National Health and Nutrition Examination Survey (NHANES III). Presented are the background papers, consensus statements, and rationale for the dietary methodologies selected for NHANES III.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control
National Center for Health Statistics

Hyattsville, Maryland
March 1992
DHHS Publication No. (PHS) 92-1464

Suggested Citation

Briefel RB, Sempos CT, eds. Dietary methodology workshop for the third National Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Stat 4(27). 1992.

Library of Congress Cataloging-in-Publication Data

Dietary methodology workshop for the third National Health and Nutrition Examination Survey (1986: Airlie, Va.)

Proceedings of the Dietary Survey Methodology Workshop for the Third National Health and Nutrition Examination Survey/sponsored by the National Center for Health Statistics.

p. cm. – (Vital and health statistics. Series 4, Documents and committee reports ; no. 27) (DHHS publication ; no. (PHS) 92-1464)

The Workshop was held in Airlie, Va. from Mar. 16 to 18, 1986.

Includes bibliographical references.

ISBN 0-8406-0441-6

1. Nutrition surveys – United States – Methodology – Congresses. 2. Diet – United States – Congresses. 3. Questionnaires – Congresses. 4. Food habits – United States – Congresses. 5. National Health and Nutrition Examination Survey (U.S.) – Congresses. I. National Center for Health Statistics (U.S.) II. Title. III. Series. IV. Series: DHHS publication ; no. (PHS) 92-1464.

[DNLM: 1. National Health and Nutrition Examination Survey (U.S.) 2. Data Collection – methods – congresses. 3. Health Surveys – United States – congresses. 4. Nutrition Surveys – United States – congresses. W2 A N148vd no. 27]

HA37.U1693 no. 27
[TX359]
362.1'0723 s – dc20
[363.8'2'0723]
DNLM/DLC
for Library of Congress

90-1338
CIP

Foreword

by Catherine E. Woteki, Ph.D., R.D., National Center for Health Statistics

How to determine what people eat has been a topic of intense research and debate in recent years. Because the methods available for use in surveys rely on the respondent's ability to recall and describe what was eaten, they are subject to errors of omission and commission that can limit the usefulness of the information for population estimates and for longitudinal studies of eating habits and later disease experience.

In 1985, when we at the National Center for Health Statistics began planning the third National Health and Nutrition Examination Survey (NHANES III), we took it as an opportunity to evaluate the methods that had been used over the preceding 25 years in our various health examination surveys. We sought the opinions of many people in the Public Health Service and other government agencies, in academic research institutions, and in industry about health and nutrition topics to be included in the survey as well as the specific methods to be used to collect information. As part of this effort, we sponsored several workshops to review, evaluate, and make recommendations about what to include in the survey. This report contains the background papers and consensus statements from a workshop on dietary survey methodology, held in March 1986, that was a fundamental part of the planning of the dietary assessment component of the survey.

The workshop participants were faced with the underlying question of choosing dietary methods that would:

- Fulfill the mission of the National Center for Health Statistics to monitor the nutritional status of the population.
- Provide baseline data on diet that would be useful in longitudinal studies of diet and health.

The workshop was organized to address statistical issues unique to dietary survey design, dietary methods most appropriate for nutrition monitoring purposes, and those most appropriate for assessing relationships with energy balance and three chronic diseases (cancer, cardiovascular disease, and osteoporosis). Many of the recommendations stemming from this workshop were incorporated into the design of NHANES III.

The papers and consensus statements developed for the workshop and contained in this volume will be useful to others who are planning surveys or to anyone who is interested in dietary assessment methods.

Acknowledgments

It is a pleasure for the cochairpersons to acknowledge the significant contributions of the authors of the papers; the discussants; Dr. Lenore Kohlmeier, who served as moderator; and the attendees to the success of the NHANES III (third National Health and Nutrition Examination Survey) Dietary Survey Methodology Workshop and to the planning process for NHANES III. We also want to acknowledge the contributions of Dr. Manning Feinleib; the Workshop Planning Committee, namely Dr. Lenore Kohlmeier, Mr. Robert Murphy, Dr. Catherine Woteki, Mr. Clifford Johnson, and Dr. Kurt Maurer; and the many staff of the National Center for Health Statistics who made the workshop possible, especially Dr. Anne Looker, Dr. Robert Kuczmarski, Dr. Deborah Winn, Dr. Stacey FitzSimmons, Mr. Dale Hitchcock, and Ms. Nancy Hamilton.

We also wish to acknowledge the Publications Branch of the National Center for Health Statistics for assistance with the final preparation of this report. Finally, special acknowledgment is given to Ms. Sharon Stewart for her patient and professional technical and editorial assistance during the preparation of this report.

Contents

Appendixes

Symbols

- - -	Data not available
. . .	Category not applicable
–	Quantity zero
0.0	Quantity more than zero but less than 0.05
Z	Quantity more than zero but less than 500 where numbers are rounded to thousands
*	Figure does not meet standard of reliability or precision

Dietary Methodology Workshop for the Third National Health and Nutrition Examination Survey

Introduction

by Ronette R. Briefel, Dr.P.H., R.D., and
Christopher T. Sempos, Ph.D., National Center for Health Statistics

Background

Planning for the third National Health and Nutrition Examination Survey (NHANES III) began in 1985 with the public solicitation of topics to be included in the survey. Later, as the study's goals and purposes were defined and target conditions selected for inclusion in the study, planning for the nutrition component necessitated the selection of a methodology or methodologies to be used for the collection of dietary intake data in NHANES III.

In the summer of 1985 initial plans were made to hold an NHANES III dietary survey methodology workshop for the purposes of reviewing, evaluating, and making recommendations regarding existing and potential dietary methodologies for NHANES III. Dietary methodologies were to be evaluated with respect to proposed target conditions for NHANES III, as well as to nutrition monitoring activities (1–3) and to the overall objectives of the survey, which include:

- To assess the health and nutritional status of the U.S. population and specific subgroups.
- To monitor changes in health and nutritional status over time.
- To provide information on the interrelationship of health and nutrition variables within population subgroups.
- To estimate the prevalence of diseases, risk factors, and health conditions and of changes over time.
- To measure met and unmet care needs related to target conditions under study.

Initial planning for the workshop involved the selection of major topics or sessions that could feasibly be covered in a 2–3 day workshop program (appendix I). The workshop planning committee (appendix II) identified six major nutrition areas of interest for the agenda:

- Statistical issues related to nutrition survey design.
- Cancer.
- Cardiovascular disease.
- Energy balance.
- Osteoporosis.
- Nutrition monitoring.

Individual scientists who could provide expertise in one or more of these areas were identified by the committee and invited to participate in the workshop (appendix III). Although it was not necessary, it was deemed desirable that invited speakers have some prior knowledge of or experience with large-scale nutrition surveys such as NHANES.

During the workshop particular emphasis was placed on the conceptual, statistical, measurement, analysis, and research issues that needed to be resolved in order to design and implement the nutrition component of NHANES III. Because NHANES III was designed to produce general descriptive statistics for the U.S. population, it was essential that the nutrition component be based upon the most current and widely accepted state-of-the-art dietary methodologies and nutrition assessment. Thus, experts in the areas of dietary methodologies, nutrition assessment, and statistics were invited to participate in a workshop designed to discuss, evaluate, and recommend dietary methodologies for inclusion in NHANES III.

Structure and scope of work

Each of the six major sessions was designed to have a working group comprised of a primary presenter, two discussants, and a National Center for Health Statistics staff liaison. An outline was provided to each working group indicating the major issues and nutrients to be considered during preparation of presentations and papers (appendixes IV and V). Issues considered by the speakers and discussants before recommending an NHANES III methodology included:

- The NHANES III survey design.
- The practicality of the dietary method in terms of use in a large survey, computer automation, the length of the interview, and cost.

- Comparability with previous methods used in NHANES (4–6) and the Hispanic Health and Nutrition Examination Survey (7).
- The 1984 recommendations of the National Academy of Sciences' Coordinating Committee on Evaluation of Food Consumption Surveys (8) with respect to linkage with the Nationwide Food Consumption Survey.
- The appropriateness of the method for different population groups.
- Interviewer training requirements.
- The ability to provide representative data of an individual's usual intake using the method.

The workshop participants were specifically charged with evaluating dietary methodologies based on the nutrition monitoring and analytic objectives of NHANES III. The primary presenter was asked to prepare an expert background paper that would be available for review by the discussants prior to the workshop. At the conclusion of the workshop, each working group was asked to prepare a consensus statement for the group's particular topic area that recommended specific dietary methodologies and strategies based upon the presentations and discussions that took place during the workshop. In addition, Dr. Lenore Kohlmeier served as the moderator and guided discussions throughout the workshop activities. Dr. Kohlmeier has extensive experience in the planning and implementation of large-scale surveys and has directed a national health and nutrition survey conducted in the Federal Republic of Germany.

Representatives from government agencies, universities, the food industry, and private scientific foundations were also invited to attend the workshop. Background materials for review were sent to all workshop participants and included an overview of the workshop objectives (appendix VI), a brief description of NHANES III (appendix VII), basic outlines used by presenters to prepare their papers (appendixes IV and V), and a dietary interviewer's manual from the Hispanic Health and Nutrition Examination Survey (9). Prior to the workshop background papers (chapter 6, appendix VIII) were provided to participants so that they would be aware of nutrition monitoring issues and past NHANES methods.

This report is organized by each of the six workshop topic areas on the program's agenda. The scientific and technical papers were prepared by the primary presenter. In addition, formal discussant papers were prepared for two of the six areas: Statistical issues (chapter 1) and nutrition monitoring (chapter 6). The consensus statements for each topic area are found in chapter 7. Chapter 8 contains Dr. Kohlmeier's remarks and commentary on the state of dietary methodologies.

After the workshop, additional nutrition planning meetings were held at the National Center for Health Statistics to take the consensus statements and recommendations into consideration with respect to the overall NHANES III planning process. Other publications by expert panels provided additional recommendations for consideration during the nutrition and dietary planning process (10,11). To place these recommendations into operation involved much planning and testing of dietary methodologies and procedures within the NHANES III framework (2,12). These survey activities and the resulting dietary procedures for NHANES III are discussed in chapter 9.

References

1. U.S. Department of Health and Human Services and U.S. Department of Agriculture. Nutrition monitoring in the United States—a progress report from the Joint Nutrition Monitoring Evaluation Committee. Washington: Public Health Service. 1986.
2. Woteki CE, Briefel RR, Kuczmarski R. Contributions of the National Center for Health Statistics. Am J Clin Nutr 47:320–8. 1988.
3. Life Sciences Research Office, Federation of American Societies for Experimental Biology. Nutrition monitoring in the United States—an update report on nutrition monitoring in the United States. Washington: Public Health Service. 1989.
4. Miller HW. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–1973. National Center for Health Statistics. Vital Health Stat 1(10a). 1978.
5. National Center for Health Statistics. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–1973. National Center for Health Statistics. Vital Health Stat 1(10b). 1977.
6. McDowell A, Engel A, Massey JT, Maurer K. Plan and operation of the second National Health and Nutrition Examination Survey, 1976–80. National Center for Health Statistics. Vital Health Stat 1(15). 1981.
7. Maurer KR. Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982–84. National Center for Health Statistics. Vital Health Stat 1(19). 1985.
8. National Academy of Sciences, National Research Council, Coordinating Committee on Evaluation of Food Consumption Surveys. National survey data on food consumption: Uses and recommendations. Washington: National Academy Press. 1984.
9. National Center for Health Statistics. Dietary interviewer's manual for the Hispanic Health and Nutrition Examination Survey, 1982–84. Intruction manual, part 15f. Hyattsville, Maryland: National Center for Health Statistics. 1986.
10. Life Sciences Research Office. Suggested measures of nutritional status and health conditions for the third National Health and Nutrition Examination Survey. Bethesda, Maryland: Federation of American Societies for Experimental Biology. 1985.
11. Life Sciences Research Office. Guidelines for use of dietary intake data. Bethesda, Maryland: Federation of American Societies for Experimental Biology. 1986.
12. Harris T, Woteki C, Briefel RR, Kleinman JC. NHANES III for older persons: Nutrition content and methodologic considerations. Am J Clin Nutr 50:1145–9. 1989.

Chapter 1
Statistical issues related to the design of dietary survey methodology for NHANES III

by Kiang Liu, Ph.D., Department of Community Health and Preventive Medicine, Northwestern University Medical School

Introduction

Several dietary factors have been evaluated for their impact on certain diseases or the risk factors of the diseases. For example, it is hypothesized that high fat intake increases, and high fiber intake decreases, the risk of colon cancer. Saturated fat and dietary cholesterol intakes have been shown to increase serum cholesterol, and polyunsaturated fat has been shown to reduce serum cholesterol. In addition, sodium intake is reported to be positively, and potassium and calcium intakes negatively, associated with blood pressure. In studies of these relationships, a major concern is the possible attenuation of the associations caused by limitations of existent dietary survey methods and improper uses of dietary data. This paper discusses the statistical issues associated with the dietary data collected by different survey methods.

Methods used in specific survey designs

Many methods are used in epidemiologic studies to collect dietary information. The dietary survey method is usually developed to fit the need of the specific study design. For a case-control study, the purpose is to recall the usual dietary pattern in the past (before the onset of the disease, it is hoped). Therefore, frequency or history methods are commonly used in this type of study (1,2). For example, Dales et al. compared the frequencies of consuming fiber-containing foods for colon cancer cases and normal controls (1), and Lubin et al. compared the average fat intakes for breast cancer patients and controls (2).

For a cross-sectional or a prospective study, the goal is to collect information on the current diet or the dietary habit in the immediate past. Thus, food records, 24-hour recalls, and frequency methods are used (3–8). In the second National Health and Nutrition Examination Survey, both the 24-hour recall and the food-frequency method were used to collect dietary information (3). Both the Framingham Study and the Western Electric study derived a quantitative method from the Burke Dietary History (4,5). The former method focused on five specific nutrients related to cardiovascular diseases. The latter was a very detailed quantitative food-frequency method based on the usual dietary intake in the past 28 days. The Hawaii study examined the relationships between several nutrients and coronary heart disease using single 24-hour recall data (6). At the second examination, 7-day food records were also used to collect the dietary information in a subsample (7). The study of Seventh-Day Adventists employed the frequency method for dietary data collection (8).

Accurate assessment of the nutrient composition of current diet is of importance in clinical trials and dietary intervention studies. Thus, multiple food records and 24-hour recalls are commonly used in these studies. For example, the National Diet-Heart Study, the Chicago Coronary Prevention Evaluation Program, and the Hypertension Control Program used 7-day food records, and the Multiple Risk Factor Intervention Trial and Lipid Research Clinic studies used single 24-hour recalls to collect dietary data (9–13). These are just a few examples of the dietary survey methods used in epidemiologic studies. Each of these methods has many different variations. For example, 24-hour recalls can be done by personal interview with three-dimensional food models or by telephone interview with or without paper food models (14,15). The frequency method is actually a family of many different methods based on the frequencies of consuming certain food items. Food intake can be assessed by personal interviews or by self-administered questionnaires (5,16–19). Some of the methods require information on portion size and the others assume a standard portion size for each food item (5,16,17). Moreover, some of the questionnaires include a very thorough detailed list of food items or groups; the others may focus only on the frequencies of the food items or groups containing certain specific nutrients (16,20).

Even though there are many dietary survey methods, from a statistical point of view, these methods can be classified into two categories:

1. Daily consumption methods—This category includes all the methods that use 1 day as a time unit, for example, single or multiple food records or 24-hour recalls.
2. Frequency methods—This category includes all the methods that ask participants to recall their usual

dietary intakes during a time period in the past—the last month, the last year, or several years ago.

Statistical issues associated with dietary survey methods

For most of the studies involving dietary surveys, the underlying assumption is that people usually maintain their dietary habits during a target period of time. For example, a case-control study of the relationship between a dietary factor or factors and a disease assumes that the disease cases maintained the dietary habits that they had before the onset of the disease. The study can then test the hypothesis that the habitual diets of the cases are higher or lower in the dietary factors of interest than those of the controls. Cross-sectional studies generally assume that the current dietary habits have been kept for a long period of time, and cohort studies further assume that the dietary habits will be continued in the future. In other words, during the time period of interest, there is no change in the mean of each of the dietary factors, and the day-to-day variation of each dietary factor is just the random variation about its mean. For most of the studies, this basic assumption is made implicitly. In some cases, the assumption may not even be reasonable because changes in dietary habits have occurred. Under this basic assumption, the two types of dietary survey methods have been used in the past to assess an individual's habitual dietary intake. The basic characteristics of these two types of methods are summarized in table 1.

Daily consumption methods

The daily consumption methods are usually based on recording or weighing the food consumed in a day or based on recalling the dietary intake in the previous 24 hours. Thus, the data are relatively more accurate than data from the other methods. However, a problem arises if an inadequate number of records or recalls are used to assess an individual's habitual intake. The diet of an individual varies from day to day. It has been shown that for most dietary factors, such as dietary cholesterol and fat intakes, the day-to-day variation within an individual (intraindividual variation) is much larger than the variation in mean intakes between individuals (interindividual variation) (21–23). Thus, the ideal way to estimate the average dietary intake of an individual is to randomly select a large number of days over the target period of time and then average the intakes of the dietary factors on these days. Unfortunately, most of the studies used the intake of only 1 day or a few days to estimate the mean intake. As a consequence, the large intraindividual variation and the inadequate number of daily measurements may attenuate the potential relationships between dietary factors and biological risk factors or diseases (21–24). As discussed later in the paper, the attenuation is more serious in analyses based on individuals and relatively less serious in analyses based on groups.

Frequency methods

The frequency methods directly estimate the average intake during the target period. Therefore, intraindividual variation is not a problem. These methods are generally easier than the collection of a large number of food records or recalls. However, the methods may suffer from the problem of inaccuracy. In some studies, a person has to report his or her usual dietary pattern during a target period of time in the remote past. Because of the limitation of a person's memory, the estimation is likely to reflect the dietary habit in the recent past. It is possible that the more remote the target period is in the past, the less accurate the estimation is. Several studies are currently ongoing to examine the ability to recall the dietary pattern in the remote past. The results of these studies should provide a better understanding of this problem.

In addition, as it is very difficult for a person to actually recall the food eaten at each meal, even for the recent past, the methods generally use questions on the usual dietary pattern. Thus, the number of meals covered by the reported survey is very unlikely to be the same as the number of meals actually taken during the target period, and the reported food is not exactly the same as the food actually eaten during the period. As a consequence, the values of each dietary factor generated from frequency methods may not be very meaningful. Several studies have reported that frequency methods tend to overestimate intakes (25,26). When the study goals rely heavily on the accuracy of the actual value of each dietary

Table 1. Characteristics of the 2 categories of dietary survey methods

Variable	*Daily consumption methods*	*Frequency methods*
Observation unit	Based on recording foods consumed in a day or recalling actual intake in previous 24 hours	Based on recalling usual dietary pattern in target period of time
Estimation method	Averaging a number of measurements to estimate the true mean intake of a dietary factor	Directly estimating the true mean intake
Accuracy of data	Relatively more accurate	Relatively less accurate
Meaningful absolute value	Yes	No
Preserving rank order	Yes, if a large number of measurements are used	Maybe
Problems caused by intraindividual variation	Yes	No
Seasonal variation taken into account	Yes, if it is designed properly (usually very difficult)	Yes, if it is designed properly
Inclusion of infrequently consumed food	Likely to miss if the number of measurements is small	Yes
Feasibility	Relatively more difficult if a large number of measurements are needed	Relatively easier
Major problem	Intraindividual variation and inadequate number of measurements	Inaccuracy of the method

factor, these methods may not be appropriate. For example, when there is a predetermined level for the deficiency of a dietary factor, the use of frequency methods is likely to have many false negative errors.

Despite these problems, these methods can be a useful tool in some cases. If the composition of the diet could be preserved (for example, the reported percentage of calories from saturated fat were the same as the true percentage), the analyses using the value of a dietary factor expressed in terms of percent of calories or per 1,000 calories would still be valid. In addition, if the rank order of each dietary factor in the population remained the same as the true rank order, the data analyzed by nonparametric techniques would generate valid results. Therefore, for a frequency method, the most critical issue is whether it can preserve the composition of each person's diet or the rank order of the population distribution of each dietary factor. When a new version of the methods is developed, it will probably be necessary to demonstrate the agreement of the rank order of each important dietary factor between the new method and some other, more accurate methods, such as an adequate number of food records or 24-hour recalls.

Effect of intraindividual-interindividual variances

Definition

For the purposes of this paper, terms are defined as follows. *True mean*—For the individual, the true mean of a dietary factor is the hypothetical average from which he or she varies during a period for which a habitual dietary pattern is maintained. *Intraindividual variation*—This is the variation of the individual's dietary intake from his or her true mean intake. *Interindividual variation*—This is the variation among the true means of individuals within a population. *Variation caused by inaccurate instrument*—This is the variation of the individual's observed mean intake from his or her true mean intake caused by the inaccuracy of the survey instrument (for example, frequency method).

Basic assumption

It is assumed that the true mean is constant during the study period for each individual; in other words, that the variation within each individual's intake represents random variation from his or her true mean and is not due to a change in habitual dietary pattern.

Daily consumption methods—Let X be an individual's true mean for a dietary factor during the period of interest. For $i = 1, \ldots, k$, let D_i be the observed value of this factor at time t_i. Because of intraindividual variation, the observed values are given by

$$D_i = X + e_i, \; i = 1, \ldots, k \qquad (1)$$

where e_i is the departure of the observed value from the true mean X because of intraindividual variation at time t_i. Let μ be the mean of X within the population and let σ_r^2 denote the variance of X (σ_r^2 being the interindividual variance). It is assumed that each individual has the same intraindividual variation and that e_i has mean 0 and variance σ_a^2 (σ_a^2 being the intraindividual variance). Furthermore, for $i \neq j$, e_i, e_j, and X are assumed to be independent random variables. As the individual's true mean for the dietary factor is not observable, X is estimated by

$$\bar{D} = \Sigma_{i=1}^{k} D_i / k$$

Frequency methods—Let X be an individual's true mean for a dietary factor during the period of interest. Let D be the estimated value of the dietary factor generated from a frequency method. Then

$$D = X + e \qquad (1)'$$

where e is the departure of the estimated value from the true mean X because of inaccurate estimation. Again, let σ_r^2 be the interindividual variance, that is, the variance of X. It is assumed that e has a mean μ_e and variance σ_e^2 (μ_e being the systematic bias caused by inaccurate estimation of the frequency method and σ_e^2 being the variance caused by inaccurate instrument). The variables X and e are assumed to be independent.

Materials and methods

Some empirical results will be given to illustrate the ideas. The data used are from two sources: Baseline 7-day food records from the Chicago Coronary Prevention Evaluation Program (CPEP) and baseline 7-day food records from the Primary Prevention of Hypertension (PPH) study. CPEP was a pilot field trial on the primary prevention of coronary heart disease by nutritional-hygienic means in 519 high-risk, coronary-prone men originally aged 40–59 years (10). The PPH study is an ongoing, randomized, controlled trial undertaken to test the ability of nonpharmacologic intervention to influence blood pressure in 200 individuals aged 30–44 years assessed to be hypertension prone (11).

Analyses of individuals

Simple correlation—Let Y be a random variable (such as a biological risk factor, say, serum cholesterol) and Y and X have a bivariate normal distribution with parameters ρ_{YX}, μ_Y, σ_Y^2, μ_X, and σ_r^2. The variable Y and e_i, $i = 1, \ldots, k$, are assumed to be independent. The correlation coefficient ρ_{YX} has to be estimated from the data. As the individual's true mean for the dietary factor is not observable, X is estimated by $\bar{D} = \Sigma_{i=1}^{k} D_i / k$. The actual correlation coefficient estimated is then

$$\rho_{Y\bar{D}} = \rho_{YX}\, c \qquad (2)$$

where

$$c^2 = \frac{1}{1 + (\sigma_a^2 / k\, \sigma_r^2)} \qquad (3)$$

or

$$c^2 = \frac{1}{1 + (\sigma_e^2 / \sigma_r^2)} \qquad (3)'$$

The constant c has different meanings for the two different types of methods. For the daily consumption methods, c is the correlation coefficient between the average value of the k measurements and the true mean value. For the frequency methods, the value c is the correlation coefficient between the estimated value (from the frequency) and the true mean value. A simple example will help to illustrate the idea. For daily consumption methods, $\sigma_a^2 / \sigma_r^2 = 3$ and the use of 1-day measurement to characterize an individual's dietary intake will result in $c = 0.5$. On the other hand, for frequency methods, $c = 0.5$ means that the correlation between the estimated value and true mean value is 0.5. In both cases, even though the sources of the attenuation of association (see details later) are different, the degrees of the attenuation are exactly the same.

For simplicity, unless it is specifically stated, the remainder of the paper will focus on the error caused by intraindividual variation (based on formula (3)). It should be understood that the impact of measurement error caused by an inaccurate instrument on various statistical analyses is exactly the same as that of intraindividual variation.

Because c is always less than 1, $\rho_{YD} < \rho_{YX}$. When the ratio σ_a^2 / σ_r^2 is very large and k is small (for example, $k = 1$), c can be much smaller than 1 and the correlation is likely to be seriously underestimated. The ratio of the intraindividual to interindividual variance can be expressed as

$$\frac{\sigma_a^2}{\sigma_r^2} = \frac{1 - \rho_{12}}{\rho_{12}}$$

where ρ_{12} is the correlation coefficient between any two measurements for the dietary factor. With use of the sample correlation coefficient r_{12} to estimate ρ_{12}, the ratio of the variances can be estimated by

$$\frac{\hat{\sigma}_a^2}{\sigma_r^2} = \frac{1 - r_{12}}{r_{12}} \qquad (4)$$

The constant c can be estimated by substituting formula (4) in formula (3). These formulas have been published previously (21). Table 2 provides the estimated ratio of the intraindividual to interindividual variances for several dietary variables. The ratios estimated from the PPH study are, in general, greater than the corresponding ratios in the CPEP study. The ratios for protein, total fat, saturated fat, polyunsaturated fat, and dietary cholesterol are all greater than 2, with the ratios for the last two being close to 5.

Based on these ratios, the possible attenuation of the association between each of these variables and risk factors can be estimated. It can be seen that, using one measurement to characterize an individual's dietary intake, except for total calories, c values are all less than 0.60. For polyunsaturated fat and dietary cholesterol, c values are about 0.40; in other words, on the average, the correlation will be reduced by about 60 percent because of the large intraindividual variation if only a 1-day food record or single recall is used to assess the individual's dietary intake.

Simple linear regression—Let Y, X, and $\bar{D}$ be defined as in "Simple correlation." Let Y be the dependent variable and X be the independent variable. Let β_1 be the regression coefficient of Y on X and β_1' be the regression coefficient of Y on $\bar{D}$. As X cannot be observed and has to be estimated by $\bar{D}$, the actual regression coefficient estimated is

$$\beta_1' = \beta_1 \cdot c^2 \qquad (5)$$

where c^2 is defined as in "Simple correlation" (27). Again, the constant c^2 can be estimated by substituting formula (4) in formula (3). Table 2 provides the estimated value of c^2 for each of the dietary factors. The values of c^2 range from 0.16 to 0.41; in other words, the percent reductions of regression coefficients range from 59 percent to 84 percent if one measurement is used in the study.

Partial correlation—Let Y, X, and $\bar{D}$ be defined as before and let $Z_1, \ldots, Z_n$ be random variables such that $Y, X, Z_1, \ldots, Z_n$ have a multivariate normal distribution.

Case 1: Assume that the controlled variables $Z_1, \ldots, Z_n$ do not have intraindividual variation. The partial correlation

$$\rho_{Y\bar{D} \cdot Z_1, \ldots, Z_n} = \rho_{YX \cdot Z_1, \ldots, Z_n} \cdot c \cdot c_1 \ldots c_n \qquad (6)$$

where for $i = 1, \ldots, n$, and $\rho_{XZ_1 \cdot Z_0} = \rho_{XZ_1}$

$$c_i^2 = \frac{1 - \rho^2_{XZ_i \cdot Z_1, \ldots, Z_{i-1}}}{1 - \rho^2_{XZ_i \cdot Z_1, \ldots, Z_{i-1}} \cdot c^2 \cdot c_1^2 \ldots c_{i-1}^2} \qquad (7)$$

and c^2 is defined as in "Simple correlation." It can be seen easily that if $\sigma_a^2 > 0$, then $c_i < 1$, for $i = 1, \ldots, n$. Thus, if the variables $Z_1, \ldots, Z_n$ do not have intraindividual variation, the partial correlation $\rho_{Y\bar{D}\, Z_1, \ldots, Z_n}$ actually estimated is always weaker than the correlation coefficient of interest, that is, $\rho_{YX \cdot Z_1, \ldots, Z_n}$. The larger the number of the controlled variables is, the larger the reduction will be. The detailed discussion of this type of attenuation was given in a paper by Liu (27).

Hypothetical example 1: Suppose that the ratio of the intraindividual to interindividual variances for dietary cholesterol is 3, and only one measurement is used to characterize an individual's intake. Thus, by formula (3), $c^2 = 0.25$. Let Y denote an individual's serum cholesterol and X and $\bar{D}$ denote the true and the observed mean dietary cholesterol intakes, respectively. Let Z denote the weight of an individual that does not vary from day to day. Assume that $\rho_{YX} = 0.6$, $\rho_{XZ} = 0.4$, and $\rho_{YZ} = 0.6$. It can be computed easily that

Table 2. Estimated ratio of intraindividual to interindividual variances

Variable	Chicago Coronary Prevention Evaluation Program				Primary Prevention of Hypertension Study			
	$\hat{\sigma}_a^2 / \sigma_r^2$	c	c^2	Number of measurements required	$\hat{\sigma}_a^2 / \sigma_r^2$	c	c^2	Number of measurements required
Total calories (kilocalories per day)	1.5	0.64	0.41	4(6)	2.0	0.57	0.32	5(9)
Protein (milligrams)	–	–	–	–	2.2	0.56	0.31	6(9)
Total fat (percent of calories)	–	–	–	–	2.6	0.53	0.28	7(11)
Saturated fat (percent of calories)	2.1	0.57	0.32	5(9)	2.7	0.52	0.27	7(12)
Polyunsaturated fat (percent of calories)	4.4	0.43	0.18	11(19)	5.2	0.40	0.16	14(22)
Dietary cholesterol (milligrams)	3.0	0.50	0.25	8(13)	4.9	0.41	0.17	13(21)
Dietary calcium (milligrams)	–	–	–	–	1.9	0.59	0.35	5(8)

NOTE: Only 1 measurement used for *c*. Number of measurements required are number needed to limit (1–*c*) to less than 15 percent and (shown in parentheses) number needed to limit (1–*c*) to less than 10 percent.

$\rho_{YX \cdot Z} = 0.491$. On the other hand, $\rho_{YD \cdot Z} = \rho_{YX \cdot Z}\, c \cdot c_1 = 0.491 \cdot 0.5 \cdot 0.935 = 0.230$. Thus, the true correlation is reduced by 53.2 percent.

Case 2: Assume that the correlated variables Y and Z and the controlled variables $Z_i, i = 1, \ldots, n$ have intraindividual variation. The impact of a large ratio of intraindividual to interindividual variances on partial correlation does not follow any pattern. For simplicity, only the first-order partial correlation coefficient will be discussed. Suppose that the variable Z cannot be observed directly and has to be estimated by $\bar{G} = \Sigma_{j=1}^{s} G_j$, where for $j = 1, \ldots, s$, $G_j = Z + h_j$. Moreover, assume that for $j = 1, \ldots, s$, h_j are independent identically distributed normal random variables with mean 0 and variance σ_{aZ}^2. Let σ_{rZ}^2 denote the variance of Z. Then,

$$\rho_{Y\bar{D} \cdot \bar{G}} = c \left(\rho_{YX \cdot Z} \sqrt{\frac{(1-\rho_{YZ}^2)(1-\rho_{XZ}^2)}{(1-\rho_{YZ}^2 d^2)(1-\rho_{XZ}^2 c^2 d^2)}} + \frac{\rho_{YZ} \cdot \rho_{XZ}(1-d^2)}{\sqrt{(1-\rho_{YZ}^2 d^2)\,(1-\rho_{XZ}^2 c^2 d^2)}} \right) \quad (8)$$

where c is defined as before and

$$d^2 = \frac{1}{1 + (\sigma_{aZ}^2 / s\, \sigma_{rZ}^2)}$$

Again, $\rho_{Y\bar{D} \cdot \bar{G}}$ can be either less than or greater than $\rho_{YX \cdot Z}$ depending on the other parameters. This property has been discussed by several researchers (27–30).

Hypothetical example 2: Let $\rho_{YX} = \rho_{YZ} = \rho_{XZ} = 0.5$, $c^2 = 0.9$, and $d^2 = 0.5$. Then $\rho_{YX \cdot Z} = 0.333 < \rho_{Y\bar{D} \cdot \bar{G}} = 0.403$. Let $\rho_{YX} = \rho_{YZ} = \rho_{XZ} = 0.5$, $c^2 = 0.25$, and $d^2 = 0.9$. Then, $\rho_{YX \cdot Z} = 0.333 > \rho_{Y\bar{D} \cdot \bar{G}} = 0.161$.

Multiple linear regression

Let Y, X, D_i, $i = 1, \ldots, k$, and Z_j, $j = 1, \ldots, n$ be defined as in "Partial correlation." Let β_1 be the partial regression coefficient of Y on X controlled for $Z_1, \ldots, Z_n$. Let β'_1 be the partial regression of Y on D controlled for $Z_1, \ldots, Z_n$. Then for $i = 1, \ldots, n$,

$$\beta'_1 = \beta_1 \cdot c^2 \cdot c_1^2 \ldots c_n^2 \quad (9)$$

where c_i^2 is defined by formula (7). Similar to the partial correlation coefficient, if $Z_1, \ldots, Z_n$ do not have intraindividual variation, the regression coefficient β'_1 actually estimated is always smaller than the regression coefficient of interest, that is, β_1. Again, the larger the number of the controlled variables is, the larger the reduction will be (28).

Hypothetical example 3: Under the same assumptions given in "Hypothetical example 1," the reduction of the regression coefficient β_1 is $100\,(1 - c^2 \cdot c_1^2) = 78$ percent. The coefficients of the other independent variables do not follow any specific pattern. The value can be either increased or decreased depending on the other parameters. Similarly, if the other independent variables also have intraindividual variation, the regression coefficient β'_1 can be either less or greater than β_1. More detailed discussion was given in the paper by Liu (27).

Multiple logistic regression (linear discriminant analysis)

The approach suggested by Walker and Duncan uses the Newton-Raphson Iteration Method to compute the maximum likelihood estimate of the logistic regression coefficients (31). Therefore, the potential attenuation cannot be clearly demonstrated. In this section, only the approach based on multivariate normal assumptions is discussed (32). Because the estimates of the coefficients are the estimates for the linear discriminant function, the results discussed in this section also apply to linear discriminant analysis. Let $X_{11}, \ldots, X_{1l}$ be the dietary and/or risk factors of the disease patients and $X_{21}, \ldots, X_{2l}$ be the dietary and/or risk factors of the nonpatients. Assume that the vectors $(X_{11}, \ldots, X_{1l})$ and $(X_{21}, \ldots, X_{2l})$ have multivariate normal distributions with mean vectors and variance-covariance matrices (U_1, Σ) and (U_2, Σ), respectively. Let $\mu = U_1 - U_2 = (\mu_1, \ldots, \mu_l)$. Then the regression coefficients

$$\beta = (\beta_l, \ldots, \beta_l) = \mu\, \Sigma^{-1}$$

Suppose that, because of intraindividual variation, X_{11} and X_{21} cannot be observed directly and have to be estimated by $\bar{D}_{11} = \Sigma_{j=1}^{k} D_{11j} / k$, and $\bar{D}_{21} = \Sigma_{j=1}^{k} D_{21j} / k$ where for $j = 1, \ldots, k$,

$D_{11j} = X_{11} + e_{11j}$ and $D_{21j} = X_{21} + e_{21j}$

Assume that e_{11j} and $e_{21j}, j = 1, \ldots, k$ are independent normal random variables with mean 0 and variance σ_a^2 and are independent of $X_{11}, \ldots, X_{1l}$ and $X_{21}, \ldots, X_{2l}$. Let β_1' denote the logistic coefficient based on D_{11} and D_{21}. Then

$$\beta_1' = \beta_1 \cdot c^2 \cdot c_2^2 \ldots c_l^2 \tag{10}$$

where c^2 is defined as in formula (3) and

$$c_2^2 = \frac{1-\rho_{12}^2}{1-\rho_{12}^2 c^2} \quad \text{and} \quad c_i^2 = \frac{1-\rho_{1i \cdot 2 \ldots i-1}^2}{1-\rho_{1i \cdot 2 \ldots i-1}^2 \cdot c^2 \cdot c_2^2 \ldots c_{i-1}^2}$$

for $i = 3, \ldots, l$ and $\rho_{1i \cdot 2 \ldots i-1}$ is the partial correlation coefficient between X_{11} and X_{1i} given $X_{12}, \ldots, X_{1i-1}$. Moreover, the odds ratio

$$e^{\beta_1'} = e^{\beta_1 \cdot c^2 \cdot c_2^2 \ldots c_l^2} \tag{11}$$

is also reduced. It can easily be seen that $c^2 < 1$ and $c_i^2 < 1$ for $i = 2, \ldots, l$. Thus, the larger the number of the variables is, the more serious the attenuation will be. The impact of measurement error on logistic regression analyses has been discussed in detail elsewhere (33).

Hypothetical example 4: Suppose that in a prospective study one wants to study the univariate relationship between the baseline dietary cholesterol intake and 12-year mortality from coronary heart disease (CHD). Assume that the true coefficient is 0.008 (that is, for every 100-milligram increase in dietary cholesterol intake, the odds ratio for CHD would increase 2.2). Assume further that the ratio of intraindividual to interindividual variances for dietary cholesterol intake is 3, and only one 24-hour recall is done to measure the dietary cholesterol intake. Thus, $c^2 = 0.25$. As a consequence of the large intraindividual variance, the coefficient reduces to $\beta_1' = 0.008 \cdot 0.25 = 0.002$ and the odds ratio for CHD corresponding to a 100-milligram increase in dietary cholesterol intake reduces to 1.22. Suppose that one wants to adjust for age in the logistic analysis. Assume that the partial regression coefficient $\beta_1 = 0.008$ and that the correlation between dietary cholesterol and age is 0.4. Then $c^2 = 0.875$ and $\beta_1' = 0.008 \cdot 0.25 \cdot 0.875 = 0.00175$. The odds ratio corresponding to a 100-milligram increase in dietary cholesterol is reduced to 1.19.

When the other variables also have intraindividual variation, β_1' does not follow any pattern. The value β_1' can be either less or greater than β_1 depending on the other parameters. Consider the case of two variables. Let X_{11} and X_{12} be the two variables for disease patients and X_{21} and X_{22} be the two variables for nondisease subjects. Suppose that X_{11}, X_{12}, X_{21}, and X_{22} cannot be observed directly and have to be estimated by $\bar{D}_{11}$, $\bar{D}_{12}$, $\bar{D}_{21}$, and $\bar{D}_{22}$, respectively. Let c^2 be the attenuation factor corresponding to $\bar{D}_{11}$ and $\bar{D}_{21}$, and d^2 be the attenuation factor corresponding to $\bar{D}_{12}$ and $\bar{D}_{22}$. Then

$$\beta_1' = \frac{c^2 [\mu_1 - (\mu_2 \rho_{12}(\sigma_1 d^2 / \sigma_2))]}{\sigma_1^2 (1 - \rho_{12}^2 \cdot c^2 \cdot d^2)}$$

and

$$\beta_1 = \frac{[\mu_1 - (\mu_2 \rho_{12}(\sigma_1 / \sigma_2))]}{\sigma_1^2 (1 - \rho_{12}^2)}$$

where σ_1 and σ_2 are the standard deviations for X_{11} and X_{12}, respectively.

Hypothetical example 5: Suppose that $\rho_{12} = 0.5$, $c^2 = 0.9$, $d^2 = 0.25$, and $\sigma_1 = \sigma_2 = 1$. If $\mu_1 = \mu_2 = 1$, $\beta_1' = 0.834 > \beta_1 = 0.667$. If $\mu_1 = 1$ and $\mu_2 = 0.5$, $\beta_1' = 0.894 < \beta_1 = 1$.

Related issues

In formulas (2) and (6), if Y also has intraindividual variations, the correlation will be further reduced. The combined impact can be measured by multiplying the attenuation factor (or factors) from Y to formulas (2) and (6).

It should be pointed out that the attenuation discussed in each analysis is the theoretical attenuation on the parameters (for example, ρ or β) caused by a large ratio of intraindividual to interindividual variances (or by the measurement error of the instrument). The estimated values (for example, r or β) are also affected greatly by sampling or other types of errors. Thus, the actual reduction in the correlation or regression coefficients may be different from the theoretical reduction.

It is important to discuss the impact of the potential attenuation on each statistical test. For simple correlation or partial correlation analysis testing whether $\rho_{YX} = 0$ or $\rho_{YX \cdot z_1 \ldots z_s} = 0$, the test statistic is

$$t = \frac{r_{Y\bar{D}}}{\sqrt{\dfrac{1 - r_{Y\bar{D}}^2}{n-2}}} \tag{12}$$

or

$$t = \frac{r_{Y\bar{D} \cdot z_1 \ldots z_s}}{\sqrt{\dfrac{1 - r_{Y\bar{D} \cdot z_1 \ldots z_s}^2}{n-s-2}}} \tag{13}$$

where n is the sample size and s is the number of covariates. Suppose that the attenuation for r is the same as the attenuation for ρ, that is, $r_{Y\bar{D}} = r_{YX} \cdot c$ or $r_{Y\bar{D} \cdot z_1 \ldots z_s} = r_{YX \cdot z_1 \ldots z_s}\, c \cdot c_1 \ldots c_s$. (In reality, because of sampling or other types of errors, this may not be the case.) Then the corresponding reduction in the t value is even larger than that for the parameters. As a consequence, the power of the test will be seriously reduced. When the sample size is very large, the denominators in formulas (12) and (13) become very small. The statistical tests still could be significant. However, no matter how large the sample size is, the correlation is still attenuated. In this case, unless ρ is very large, a very small but significant estimated value is likely to be obtained.

For simple and multiple linear regression analyses, because the test for H_o: $\beta = 0$ is equivalent to the test for H_o:$\rho = 0$, the above discussion can also be applied to regression analyses.

For multiple logistic regression,

$$\text{Var}(\hat{\beta}_1) = \hat{\sigma}_v^{11}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$$

where $\hat{\beta}_1$ is the estimated value for β_1; σ_v^{11} is the element in the first row and the first column of the inverse matrix of the variance-covariance matrix V for the vector $(\bar{D}_{11}, X_{12}, \ldots X_{1l})$ and $\hat{\sigma}_v^{11}$ the corresponding estimate; N_1 and N_2 are the number of patients and nonpatients, respectively. The test statistic is

$$z = \frac{\hat{\beta}_1'}{\sqrt{\hat{\sigma}_v^{11}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

It can be shown that the potential attenuation for the z value is $c \cdot c_2 \ldots c_l$. If β_1 is very large and both N_1 and N_2 are very large, the test still can be significant. However, the regression coefficient remains attenuated. In reality, the number of patients is likely to be small; thus, the potential attenuation of the test statistic is likely to be large.

It is possible to estimate the potential attenuation for simple correlation and regression analyses and to estimate the number of measurements needed.

For correlation, one can estimate the number of measurements by the formula

$$k = \frac{c^2}{(1-c^2)} \cdot \frac{\hat{\sigma}_a^2}{\sigma_r^2} \qquad (14)$$

where $(1-c)100$ is the predetermined amount of attenuation.

Table 2 provides the needed number of measurements for $c \geq 0.85$ and for $c \geq 0.90$ based on the CPEP and PPH data. For most dietary factors, 7 to 14 daily food records are needed for $c > 0.85$.

For simple linear regression, the number of measurements can be estimated by

$$k = \frac{c}{(1-c)} \cdot \frac{\hat{\sigma}_a^2}{\sigma_r^2}$$

where $(1-c)100$ is the predetermined amount of attenuation.

Analyses of groups

Analysis of variance—A large ratio of intraindividual to interindividual variances can seriously reduce the power of the statistical test in analyses of variance. For $i = 1, \ldots, s$, and $j = 1, \ldots, n_i$, let X_{ij} be independent normal random variables with mean μ_i and variance σ_r^2. Suppose that X_{ij} cannot be observed and has to be estimated by $\bar{D}_{ij\cdot j}$, where $\bar{D}_{ij\cdot} = \Sigma_{w=1}^{k} D_{ijw} / k$, the average of k observed values for X_{ij}. Again, let σ_a^2 denote the intraindividual variance. Under a set of the assumptions of normality and independence, the power for testing $H_o{:}\tau_i = \mu_i - \mu_{\cdot} = 0, i = 1, \ldots, s$ against $H_a{:}\tau_i \neq 0$, for some i, is

$$\text{Power} = P\{F^* > F(1-\alpha; s-1, \Sigma_{i=1}^{s} n_i - s) \mid \phi\}$$

where F^* has a noncentral F distribution with noncentrality parameter

$$\phi = \frac{1}{\sigma_r}\sqrt{\frac{\Sigma_{i=1}^{s} n_i \tau_i^2}{s}}$$

and degrees of freedom $s-1$ and $\Sigma_{i=1}^{s} n_i - s$ (34).

Because X_{ij} cannot be observed directly and has to be estimated by $\bar{D}_{ij\cdot}$, the noncentrality parameter is

$$\phi' = \phi \cdot c$$

As $\phi' < \phi$, the power is reduced.

Hypothetical example 6: Suppose that $s = 2$, $\Sigma_{i=1}^{s} n_i - s = 10$, $\alpha = 0.05$, $\phi = 3$, and $c = 0.5$. The power corresponding to $\phi = 3$ is 0.983 (34). The power corresponding to $\phi = 1.5$ is 0.50.

Estimation—A large ratio of intraindividual to interindividual variances can also have an impact on estimation of parameters. Let X_j and $\bar{D}_{\cdot j}$ be defined as usual. Again, assume a set of standard assumptions of normality and independence. Based on X_j, the $100 \cdot p$ percent confidence interval for the mean μ is

$$\bar{X} \pm Z_{(1+p)/2}\sqrt{\sigma_r^2 / n}$$

Based on $D_{\cdot j}$, the confidence interval for μ is

$$\bar{D}_{\cdot\cdot} \pm Z_{(1+p)/2}\sqrt{\sigma_r^2 / n \cdot c^2}$$

where $Z_{(1+p)/2}$ is the $100(1+p)/2$th percentile of the standard normal distribution.

Because $c < 1$, the length of the confidence interval is increased. However, if the sample size n is very large, the length of the interval will be very small, even with the factor $1/c$.

Misclassification—Many statistical problems are caused by misclassification. If the variable that has intraindividual variation is used for classifying individuals into different categories, a proportion of the individuals will be misclassified. Based on the same assumptions given in "Simple correlation," the probability of misclassifying an individual whose true mean intake is in the lower $100{\cdot}p$ percent of the distribution of X into the upper $100{\cdot}p$ percent of the distribution based on the observed value $\bar{D}$ is

$$P(Z_2 > Z_{1-p} \mid Z_1 < Z_p)$$

where Z_1 and Z_2 are two standard normal random variables having a joint bivariate normal distribution with correlation coefficient

$$\rho_{Z_1 Z_2} = c$$

and Z_p is the $100{\cdot}p$th percentile of the standard normal distribution.

Hypothetical example 7: Individuals are classified into the low and the high level of dietary cholesterol intake based on one measurement. Suppose that $c = 0.5$. Then $P(Z_2 > 0 \mid Z_1 < 0) = 0.333$. That is to say, 33.3 percent of the individuals in the low level are misclassified into the high level and 33.3 percent of the individuals in the high level are misclassified into the low level.

Measure of risk in cross-sectional study—Both the risk ratio (relative risk) and odds ratio can be attenuated by misclassifying individuals into wrong categories. For $i, j = 1, 2$, let P_{ij} be the probability that an individual is classified into the ith category of the dietary factor and the jth category of the disease status, and $P_{.1}$, $P_{.2}$, $P_{1.}$, and $P_{2.}$ be the marginal probabilities, where $i = 1$ if high group and $i = 2$ if low group; $j = 1$ if disease and $j = 2$ if no disease. Let P_L be the conditional probability that an individual is misclassified to the high group given that the person is in the low group and let P_H be the conditional probability that an individual is misclassified to low given that the person is in high. Table 3 provides the expected number in each cell with and without misclassification. Without misclassification, the risk ratio R and odds ratio O are

$$R = \frac{P_{11} / P_{1.}}{P_{21} / P_{2.}}$$

and

$$O = \frac{P_{11} P_{22}}{P_{21} P_{12}}$$

With misclassification, the risk ratio R_M and odds ratio O_M change to

$$R_M = \frac{[P_{11} (1-P_H) + P_{21} P_L] / P_{1.}}{[P_{11} P_H + P_{21} (1-P_L)] / P_{2.}}$$

and

$$O_M = \frac{[P_{11} (1-P_H) + P_{21} P_L] \, [P_{12} P_H + P_{22} (1-P_L)]}{[P_{11} P_H + P_{21} (1-P_L)] \, [P_{12} (1-P_H) + P_{22} P_L]}$$

It can be seen easily that if $R > 1$, $R_M < R$ and $O_M < 0$.

Hypothetical example 8: Consider table 4, which deals with probabilities. Suppose that $c = 0.5$. Based on formulas given in "Cross-sectional study," $P_L = 0.1418$ and $P_H = 0.5674$. Then $R_M = 1.25$, $O_M = 1.40$, $R = 2$, and $O = 3$.

Measure of risk in cohort study—Let N_1 and N_2 be the numbers of individuals in the high and low groups at baseline, respectively. Let P_1 and P_2 be the probabilities of developing the disease in the high and low groups after t years of followup. Moreover, let P_L and P_H be defined as before. Then $R = P_1 / P_2$, $O = P_1(1-P_2) / P_2(1-P_1)$,

$$R_M = \frac{[N_1 P_1 (1-P_H) + N_2 P_2 P_L] / N_1}{[N_1 P_1 P_H + N_2 P_2 (1-P_L)] / N_2} \tag{15}$$

and

$$O_M = \frac{[N_1 P_1 (1-P_H) + N_2 P_2 P_L] \, [N_1 (1-P_1) P_H + N_2 (1-P_2)(1-P_L)]}{[N_1 P_1 P_H + N_2 P_2 (1-P_L)] \, [N_1 (1-P_1)(1-P_H) + N_2 (1-P_2) P_L]} \tag{16}$$

It should be pointed out that if, at the tth year, only a random subsample of the low group is used for calculation of risk ratio or odds ratio, these expected ratios can still be computed by replacing the new subsample size for N_2.

Hypothetical example 9: Let $N_1 = 200$ and $N_2 = 800$. Suppose that $P_1 = 0.30$ and $P_2 = 0.10$. Furthermore, suppose that $P_L = 0.1418$ and $P_H = 0.5674$. Then $R = 3$, $O = 3.86$, $R_M = 1.45$, and $O_M = 1.56$.

Measure of risk in case-control study—Let P_1 and P_2 be the proportion of the high level of a dietary factor (without intraindividual variation) in the case and the control groups, respectively. Let N_1 and N_2 be the number of people in the case and the control groups and let P_H and P_L be the probabilities of misclassification defined previously. Then

$$O = \frac{P_1 (1-P_2)}{(1-P_1) P_2}$$

and

$$O_M = \frac{[P_1 (1-P_H) + (1-P_1) P_L] \, [P_2 P_H + (1-P_2)(1-P_L)]}{[P_1 P_H + (1-P_1)(1-P_L)] \, [P_2 (1-P_H) + (1-P_2) P_L]}$$

For most of the case-control studies, the dietary information is collected by the frequency methods. Thus, the misclassification is not due to intraindividual variation but is due to the errors made by the direct estimation of an individual's true mean intake of the frequency methods.

Hypothetical example 10: Let $P_1 = 0.3$ and $P_2 = 0.1$ and let $P_H = 0.5674$ and $P_L = 0.1418$. Then $O = 3.86$ and $O_M = 1.44$.

Related issues—For analysis of variance and estimation of means, the impact of a large ratio of intraindividual

Table 3. Expected number in each cell without and with misclassification

Dietary factor	Total	Disease: Yes	Disease: No
		Without misclassification	
Total	N	$N P_{1.}$	$N P_{.2}$
High	$N P_{1.}$	$N P_{11}$	$N P_{12}$
Low	$N P_{2.}$	$N P_{21}$	$N P_{22}$
		With misclassification	
Total	N	$N P_{.1}$	$N P_{.2}$
High	$N P_{1.}$	$N [P_{11} (1-P_H) + P_{21} P_L]$	$N [P_{12} (1-P_H) + P_{22} P_L]$
Low	$N P_{2.}$	$N [P_{11} P_H + P_{21} (1-P_L)]$	$N [P_{12} P_H + P_{22} (1-P_L)]$

Table 4. Probability of cross-classification into the cells

Dietary factor	Disease: Yes	Disease: No	Total
High	0.1	0.1	0.2
Low	0.2	0.6	0.8
Total	0.3	0.7	1.0

to interindividual variances is generally not very severe if the sample size in each group is very large.

For a classification variable, a large ratio of intraindividual to interindividual variances and an inadequate number of measurements will lead to a large proportion of misclassification. The impact can be very serious. For example, the reduction in risk ratio and odds ratio can be very large, and the attenuation will not be changed by increasing the sample size.

In the analysis of variance or covariance, misclassification between groups will reduce the difference between the group means and increase the variance. Thus, the power will be reduced. In some situations, the assumption of the underlying distribution may become invalid after misclassification.

Practical problems

Several practical problems related to intraindividual variation need to be considered in designing a study to measure dietary intake.

Feasibility of multiple records—The number of food records or 24-hour recalls needed to characterize accurately an individual's dietary intake may create practical difficulties. For example, in a large-scale study, collection of 7- to 14-day food records requires a large amount of time and effort from trained nutritionists and participants and thus may not be feasible. This is a common problem faced in such studies. Unfortunately, there is no perfect solution to this problem.

The following steps may help to partially reduce the severity of the problem.

1. Reduction of intraindividual variation caused by other factors: Usually food records and recalls are converted into nutrient codes by trained coders. However, in many cases the coding requires a subjective judgment from the coders, and coders may not always follow the coding rules. Thus, a part of the observed intraindividual variation of an individual's dietary intake may be due to intracoder and intercoder variation in coding food records or 24-hour recalls. Standardization of coders and their ongoing training and monitoring for consistency in coding food records or recalls could reduce the observed intraindividual variation.

   For 24-hour recalls, intrainterviewer or interinterviewer variations in interviewing techniques can also contribute to intraindividual variation. However, interviewing techniques can be standardized by training of the interviewers. For example, Beaton et al. found that well-trained interviewers contributed very little to the intraindividual variation (22).
2. Estimation of the ratio of intraindividual to interindividual variances for the study population: The number of measurements required may vary from population to population. For example, the ratio of the intraindividual to interindividual variances for the 24-hour urine sodium excretion estimated based on data in Chicago was about 3 (35,36). However, the ratios estimated based on data in Belgium and China were less than 2 (37,38). Therefore, the number of measurements needed to characterize an individual's sodium intake is probably less for the studies in Belgium and China than the study for Chicago. Similarly, for different populations the number of measurements needed to characterize an individual's dietary intake may not always be the same. It may be worthwhile to estimate the required number of measurements for the study population. However, for a study population that includes different subpopulations, such as the third National Health and Nutrition Examination Survey (NHANES III), it may complicate the data analyses and sometimes even create biases if different numbers of measurements are used for different subpopulations. Thus, multiple numbers of measurements are not recommended for a study even though the study population includes many different subpopulations.
3. Selection of an optimal number of measurements: It can be seen in table 2 that when the reduction in correlation is not very large (that is, c is close to 1), even a small decrease in reduction requires a large number of measurements. For example, for total calories, five measurements decrease the reduction in correlation from 43 percent to 15 percent; an additional four measurements decrease the reduction only from 15 percent to 10 percent. Therefore, one should select an appropriate number of food records or recalls to best fit the study.

Random versus consecutive measurements—Another problem that needs to be considered is the choice between randomly scheduled food records or recalls and consecutive daily records. From a statistical point of view, if the number of measurements is large, randomly scheduled food records or recalls during the study period are certainly more appropriate. However, the number of measurements is usually not very large. If for a majority of people there is a habitual pattern of dietary intakes for different days in a week, randomly scheduled records or recalls may not always be the best method. Hankin, Reynolds, and Margen suggested that dietary intake during weekend days may be slightly different from the intake during weekdays (39). If this is the case, the difference in dietary intakes between weekend days and weekdays should also be taken into account in the random scheduling of the measurements. For example, if seven measurements are needed, five of them can be randomly scheduled on weekdays and two on weekend days.

Consecutive food records may be easier to do than randomly scheduled records or recalls. If the day-to-day variation of an individual's dietary intakes on consecutive days is not independent (that is, the variables e_i, $i = 1, \ldots, k$ in "Simple correlation" are not independent), use of consecutive food records may be inappropriate. However, this is probably not the case. For example, in the study done by Hankin, Reynolds, and Margen (39), the correlation coefficient of each nutrient between two consecutive days was not any higher than that between two nonconsecutive days. Thus, even though use of consecutive food records in recalls is not as good as randomly scheduled records or recalls, it may still be a reasonable alternative when random food records or recalls are not feasible.

Seasonal variation — An individual's dietary intakes may vary from season to season. From a statistical point of view, seasonal variation is not a random variation but rather a change in mean intake. If it is large, it may create some biases in the study. Sempos et al. conducted a 2-year dietary survey of middle-aged women (40). The participants were asked to record their diet on 2 randomly selected days per month over a 2-year period. The results indicated that the intake of certain food groups varied from season to season. On the other hand, Potter, McMichael, and Bonett reported that there was no evidence of seasonal variation (41). Because of rapid development in modern technologies, almost all foods are available year round. It is possible that seasonal variation is not as large as before. Unfortunately, the available literature does not provide adequate information to evaluate how large the seasonal variation is. The diet data collected by several investigators can be further analyzed to answer this question (17,40).

If the seasonal variation is very large, it should be taken into account in the study design to avoid any possible biases. Ideally, multiple records or recalls should be scheduled for each participant for each season to account for seasonal variation. Unfortunately, for a large-scale study, the collection of multiple records or recalls can be very difficult. A carefully designed study can help to partially solve the problem. For example, for the comparison of the average intakes between different subgroups (such as, age, sex, and race), the possible biases caused by seasonal variation can be reduced if the diet surveys for each subgroup are randomly scheduled throughout the year. For many biological risk factors, the level is stabilized within a relatively short period of time after a diet change. For example, serum cholesterol levels are stabilized 3 to 4 weeks after changing from a high-fat diet to a low-fat diet (42). Thus, for studying the relationships between dietary factors and biological risk factors, measuring the biological risk factor on the same day of the dietary survey may help to prevent the potential attenuation caused by seasonal variation.

Statistical correction — For simple correlation or linear regression analysis, it is possible to correct statistically for the attenuation caused by a large intraindividual variation. However, the correction is not recommended unless the sample size is very large and the estimated value of the ratio of intraindividual to interindividual variances is very accurate. When the sample size is small, the observed correlation or regression coefficient and the estimated ratio of intraindividual to interindividual variances may also be greatly affected by sampling variations. Thus, statistical correction of an observed coefficient may produce misleading results. Moreover, statistical correction is not appropriate for the other type of analyses, such as partial correlation, multiple linear regression, multiple logistic regression, analysis of variance, and risk ratio analyses.

Recommendations

Should multiple dietary records be collected in NHANES III?

From a statistical point of view, it is more appropriate for NHANES III to use as its primary method the daily consumption method rather than the frequency method. More specifically, the recommended method is three 24-hour recalls, with two randomly assigned for weekdays and one for weekend days. If the three 24-hour recalls using person-to-person interview are not feasible, the first 24-hour recall should be done by person-to-person interview using food models, and the subsequent two should be done by telephone interview using paper models. This is recommended for the following reasons.

1. One of the major purposes of NHANES III is to estimate the average dietary intakes of different nutrients for all subgroups (age, sex, race, socioeconomic, regional, and so forth). With a large sample size in each of the subgroups, the estimation can be reasonably accurate even with one measurement. On the other hand, the frequency methods cannot provide accurate absolute value for each nutrient. It is very likely to overestimate or underestimate the average value. Therefore, the daily consumption method should be used.
2. In the two previous NHANES surveys, a single 24-hour recall was used as the primary survey method. In order to compare the secular trend of the dietary intake for each dietary factor, the same survey method should be used. Otherwise, biases will be created by the use of different methods.
3. Food records or 24-hour recalls provide the raw data on the actual foods eaten by an individual. However, frequency methods usually group different foods together. As a consequence, a great deal of information may be lost. For example, many years from now some nutrients or trace elements may become important. It is possible to estimate the baseline intake for each individual if the raw data are available. The information cannot be retrieved if certain foods are grouped together.
4. If an inadequate number of 24-hour recalls is used to characterize an individual's dietary intake, the association between the dietary variable and the biological risk factor will be attenuated by the intraindividual variation. On the other hand, if a frequency method is used to assess an individual's dietary intake, the association will also be attenuated by the inaccuracy of the method. The attenuation corresponding to most of the frequency methods is about the same as the attenuation corresponding to three 24-hour recalls.

The use of two 24-hour recalls done by telephone interview and one 24-hour recall done by person-to-person interview seems to be a possible alternative if three 24-hour recalls done person to person are not feasible. Two studies had encouraging results on the accuracy of the telephone interview 24-hour recalls (14,15). This method should be further evaluated before use in NHANES III.

Limitations of daily consumption method

The major limitation of the daily consumption method is the possible attenuation caused by intraindividual variation.The details have been discussed previously. For group analyses, if the sample size is large, the impact is not very serious. For analyses of individuals and risk analyses, the impact can be very serious. If the association between a dietary factor and a biological variable is strong, the relationship may be attenuated but remains significant. However, if the association is weak, the relationship will be attenuated and may become unobservable.

The other limitation is the possible omission of infrequently consumed foods, such as liver and alcohol. These 24-hour recalls may not be adequate to cover the intake of these foods. The addition of a frequency questionnaire by Willett et al. or Block et al. should be seriously considered (17,43).

It should be pointed out that there is no perfect dietary survey method. Every method has some limitations. Considering the statistical problems and feasibility of each method the most reasonable choice as the primary survey method for NHANES III is three 24-hour recalls.

References

1. Dales LG, Friedman GD, Ury HK, et al. A case-control study of relationships of diet and other traits to colorectal cancer in American blacks. Am J Epidemiol 109:132–44. 1978.
2. Lubin JH, Burns PE, Blot WJ, et al. Dietary factors and breast cancer risk. Int J Cancer 28:685–9. 1981.
3. McDowell A, Engel A, Massey JT, Maurer K. Plan and operation of the second National Health and Nutrition Examination Survey, 1976–80. National Center for Health Statistics. Vital Health Stat 1(15). 1981.
4. Mann GV, Pearson G, Gordon T, Dawber TR. Diet and cardiovascular disease in the Framingham Study: I. Measurement of dietary intake. Am J Clin Nutr 11:200–25. 1962.
5. Shekelle RB, Stamler J, Paul O, et al. Dietary lipids and serum cholesterol level: Change in diet confounds the cross sectional association. Am J Epidemiol 115:506–14. 1982.
6. Tillotson JL, Kato MA, Nichaman MZ, et al. Epidemiology of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: Methodology for comparison of diet. Am J Clin Nutr 26:177–84. 1973.
7. Kato MA, Tillotson JL, Nichaman MZ, et al. Epidemiological studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California. Am J Epidemiol 97:372–85. 1973.
8. Kahn HA, Phillips RL, Snowdon DA, Choi W. Association between reported diet and all-cause mortality: Twenty-one-year follow-up on 27,530 adult Seventh-Day Adventists. Am J Epidemiol 119:775–87. 1984.
9. National Diet-Heart Study Research Group: The National Diet-Heart Study final report. Circulation 37 (suppl 1):1–428. 1968.
10. Farinaro E, Stamler J, Upton M, et al. Plasma glucose levels: Long-term effect of diet in the Chicago Coronary Prevention Evaluation Program. Ann Int Med 86:147–54. 1977.
11. Stamler R, Grimm R, Gosch FC, et al. Control of high blood pressure by nutritional therapy: Final report of a 4-year randomized controlled trial—The Hypertension Control Program. JAMA 257:1484–91. 1987.
12. Caggiula AW, Christakis G, Farrand M, et al. The Multiple Risk Factor Intervention Trial (MRFIT). IV. Intervention on blood lipids. Prev Med 10:443–75. 1981.
13. Lipid Research Clinics Program. The Lipid Research Clinics Coronary Primary Prevention Trial results: I. Reduction in incidences of coronary heart disease. JAMA 251:351–64. 1984.
14. Krantzler NJ, Mullen BJ, Schutz HG, et al. Validity of telephoned diet recalls and records for assessment of individual food intake. Am J Clin Nutr 36:1234–42. 1982.
15. Posner BM, Borman CL, Morgan JL, et al. The validity of a telephone-administered 24-hour dietary recall methodology. Am J Clin Nutr 36:546–53. 1982.
16. Mahalko JR, Johnson LK, Gallagher SK, Milne DB. Comparison of dietary histories and seven-day food records in a nutritional assessment of older adults. Am J Clin Nutr 42:542–53. 1985.
17. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65. 1985.
18. Chu SY, Kolonel LN, Hankin JH, Lee J. A comparison of frequency and quantitative dietary methods for epidemiologic studies of diet and disease. Am J Epidemiol 119:323–34. 1984.
19. Hankin JH, Messinger HB, Stallones RA. A short dietary method for epidemiologic studies. III. Development of questionnaire. Am J Epidemiol 87:285–98. 1967.
20. Samet JM, Humble CG, Skipper BE. Alternatives in the collection and analysis of food frequency interview data. Am J Epidemiol 120:572–81. 1984.
21. Liu K, Stamler J, Dyer A, et al. Statistical methods to assess and minimize the role of intra-individual variability in observing the relationship between dietary lipids and serum cholesterol. J Chron Dis 31:399–418. 1978.
22. Beaton GH, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J Clin Nutr 32:2546–59. 1979.
23. Sempos CT, Johnson NE, Smith EL, Gilligan C. Effects of intra-individual and inter-individual variation in repeated diet records. Am J Epidemiol 121:120–30. 1985.
24. Jacobs DR, Anderson JT, Blackburn H. Diet and serum cholesterol: Do zero correlations negate the relationship? Am J Epidemiol 110:77–87. 1979.
25. Young CM, Chalmers FW, Church HN, et al. A comparison of dietary study methods. J Am Diet Assoc 28:124–8. 1953.
26. Block B. A review of validations of dietary assessment methods. Am J Epidemiol 115:492–505. 1982.
27. Liu K. Measurement error and its impact on partial correlation and multiple linear regression analyses. Am J Epidemiol 127:864–74. 1988.
28. Snedecor GW, Cochran WG. Statistical methods. Ames, Iowa:Iowa State University Press. 1974.

29. Fleiss JL, Shrout PE. The effects of measurement errors on some multivariate procedures. Am J Public Health 67:1188–91. 1977.
30. Kupper LL. Effect of the use of unreliable surrogate variables on the validity of epidemiologic research studies. Am J Epidemiol 120:643–8. 1984.
31. Walker SH, Duncan DB. Estimation of the probability of an event as a function of several independent variables. Biometrika 54:167–79. 1967.
32. Truett J, Cornfield J, Kannel W. A multivariate analysis of the risk of coronary heart disease in Framingham. J Chron Dis 20:511–24. 1967.
33. Liu K, Sempos C. Measurement error and its impact on multiple logistic regression analysis. Unpublished manuscript.
34. Neter J, Wasserman W. Applied linear statistical model. Homewood, Illinois: Richard D. Irwin, Inc. 1975.
35. Liu K, Cooper R, McKeever J, et al. Assessment of the association between habitual salt intake and high blood pressure: Methodological problem. Am J Epidemiol 110:219–26. 1979.
36. Liu K, Cooper R, Soltero I, Stamler J. Variability in 24-hour urinary sodium excretion in children. Hypertension 1:631–6. 1979.
37. Joossens JV, Claessens J, Geboers J, Claes JH. Electrolytes and creatinine in multiple 24-hour urine collections (1970–1974). In: Epidemiology of arterial blood pressure, Kesteloot H, Joossens J, eds. The Hague: Martinus Nijhoff Publishers bv.
38. Hsiao CK, Wang SY, Hong ZG, et al. Timed overnight sodium and potassium excretion and blood pressure in steel workers. J Hypertens 4:345–50. 1986.
39. Hankin JH, Reynolds WE, Margen S. A short dietary method for epidemiologic studies: II. Variability of measured nutrient intakes. Am J Clin Nutr 20:934–45. 1967.
40. Sempos CT, Johnson N, Smith EL, Gilligan C. A two-year dietary survey of middle-aged women: Repeated dietary records as a measure of usual intakes. J Am Diet Assoc 84:1008–13. 1985.
41. Potter JD, McMichael AJ, Bonett AZ. Diet, alcohol and large bowel cancer; a case-control study. Proc Nutr Soc 7:123–6. 1982.
42. Connor WE, Hodges RD, Bleiler RE. Effect of dietary cholesterol upon serum lipids in man. J Lab Clin Med 57:331–42. 1961.
43. Block G, Hartman AM, Dresser CM, et al. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 124:453–69. 1986.

First discussant's remarks about statistical issues

by Harold A. Kahn, M.A., Consultant in epidemiology

My one serious disagreement with Dr. Liu's paper is in my unwillingness to accept the validity of the 24-hour dietary recall as established. I do not think that observation of communal meals, comparison of 24-hour recalls with diet histories, frequency counts, or other diet records that themselves have not been validated or the finding of statistically significant correlations constitute evidence of validity.

Because it is reasonable to assume that persons who cannot remember what they ate yesterday are unlikely to be able to accurately average their diet over the past year in order to report their usual intake, the situation regarding validity logically falls under one of the three following categories:

- All methods have been validated.
- Only daily methods (24-hour recall or food record) have been validated.
- No methods have been validated

If all methods are valid, because of simplicity and ease of standardization, the frequency methods would be the obvious choice.

If only daily methods are valid, it would be appropriate to use them to compute the average diet for groups. Individual classification by diet could also be attempted by use of multiple days of data using the adjustments specified by Dr. Liu. An alternative that permits unbiased estimation of the odds ratio relating diet to disease after adjustment for multiple covariables is outlined below, modified from the methods described in the National Academy of Sciences, National Research Council monograph *Nutrient Adequacy* (1).

1. Transform the dietary data to approximate normality.
2. Classify all persons by disease outcome (for example, sick or well) or by a dichotomous physiologic variable (for example, serum cholesterol less than 240 milligrams per deciliter or 240 milligrams per deciliter or more).
3. Collect two random independent 24-hour dietary recalls (or records) from each subject and for each "disease" group estimate the mean ($\bar{x}$) and variance of individual dietary intakes ($=s_b^2$).
4. From the paired values for all individuals estimate the variance within individuals ($=s_w^2$, assuming the data are poolable).
5. Subtract $s_w^2/2$ from s_b^2 to estimate the true variance between individuals, free from the additional variance of random daily departures from the individual's true mean ($= s_b^2$, revised).
6. Use the $\bar{x}$ together with s_b (revised) to estimate the proportion of persons with intake above some cutpoint.
7. Say this is 0.05 for the "sick" and 0.03 for the "well," and for convenience assume that the number of sick and well persons is 100 each.

We can now make a 2 × 2 table:

	Sick	Well
Risk factor +	5	3
Risk factor −	95	97
	100	100

$$\text{Estimated odds ratio} = \frac{5 \times 97}{3 \times 95} = 1.70$$

8. Carry out the above for each stratum of covariables for which adjustment is desired (for example, young men with high serum cholesterol, young men with low serum cholesterol).
9. Use the Mantel-Haenszel procedure to estimate an adjusted odds ratio free of bias caused by variance within individuals.

The last possibility, that no methods have been validated, suggests the urgent need for a validity test, perhaps using the physiologic fact that for normal adults in our society a 24-hour urine specimen reflects about 90 percent of the dietary intake of protein, riboflavin, and thiamine. Because it is not today's intake that is reflected in today's urine specimen but a weighted average of the intake over the previous 3–4 days, it is not correct to match urine with daily diet data for each individual; however, it is possible to match group averages. Obtain volunteers who agree to be contacted on a random date during a study period of about 4 months to provide a 24-hour recall interview and begin a 24-hour urine collection on that date. As the 24-hour recalls are collected on a random date, it is possible to use them to estimate the group average intake during the study period according to the recall method. Similarly, because the few days preceding a random date are also randomly chosen, the analysis of 24-hour urine specimens for total nitrogen, riboflavin, and thiamine can be used to estimate the group average intake for protein, riboflavin, and thiamine during the study period according to the urinalysis. If 24-hour specimens are, in fact, collected and if the averages differ by more than can be expected from random sampling error, it would be evidence against the 24-hour recall procedure. In addition, rank correlation coefficients can be compared (adjusted as in Dr. Liu's paper). Failure to find evidence against the recall procedure using these nutrients does not prove the case for other nutrients. However, this is the traditional way science progresses, by attempting to reject hypotheses. Because these nutrients are pervasive in the diet, "passing" this validity test would certainly strengthen the case for using the 24-hour recall method. In addition, "passing" the test for group means would justify using multiple 24-hour recalls for more precise estimates of individual diets.

One other suggestion (if no methods are known to be valid) is to analyze frequency data only for the "never eat" and "eat daily" categories. I want to make a final plea for stressing prospective studies. While we try to determine how diet affects disease, let us not forget that disease affects diet. Asking someone to report what he ate "before getting sick" or "2 years ago," when it is unclear that he can tell us what he ate yesterday, requires an unjustified leap of faith.

Reference

1. National Academy of Sciences, National Research Council, Subcommittee on Criteria for Dietary Evaluation, Coordinating Committee on Evaluation of Food Consumption Surveys. Nutrient adequacy: Assessment using food consumption surveys. Washington: National Academy Press. 1986.

Second discussant's remarks about statistical issues in design

by George H. Beaton, Ph.D., Department of Nutritional Sciences, University of Toronto

Initial remarks

The paper by Dr. Kiang Liu marks a major contribution to the evolution of thinking about design issues related to the collection of food intake. But then we have come to expect his continuing contribution in this area; he has been leading the way for a decade or more.

Today we know enough of this story to make it inexcusable to ignore it in designing large studies such as NHANES III. The specific moral that we must address is that one must carefully define the questions before designing the study. It seems clear that there is no perfect dietary methodology for inclusion in NHANES III. There is no methodology that will provide data suitable for analysis of all of the types of questions that users have attempted to put to previous NHANES reports.

I think we are in a position to begin building a matrix of types of questions, analytical approaches, and dietary methodology requirements—a matrix that would permit a sensible choice of methods together with a clear declaration of what can and cannot be done with NHANES III data. It is not my role to try to determine which types of questions should have priority, but I will try to begin building the matrix linking question and design.

Technical comments on paper

First, at a level of supplementary information, table 2 of Dr. Liu's paper (chapter 1) presents some estimates of the ratio of intraindividual to interindividual variances. These were based on data from two studies in which he is involved. Recently a National Academy of Sciences (NAS) report drew together estimates published in the literature (1). These are compared in table 1 of this section. The point to note is that, in some cases, the ratios appear to be appreciably different across studies; in other cases, the ratios are similar.

Two important messages arise: (a) The phenomenon that Dr. Liu discusses is one that has been found in every data set examined to date and must be recognized in design and analysis; (b) there is strong indication, as might be expected, that the variance ratios differ between studies, perhaps by age, sex, and cultural group and with the particular methodology applied. The latter is important, as the intraindividual variation includes not only true biological variation within the individual but also random methodologic variation. The interindividual variation will reflect both the characteristics of the particular population group and also sampling effects. These ratios must be established internally within a data set if they are to be used in analysis or in interpreting analyses.

Second, Dr. Liu raises the question of consecutive versus independent days in increasing the number of observations. Recently Morgan, Johnson, and Goungetas examined this effect, and they found that it is a persistent effect and appears to vary with the level of intake (2). They urge a more extensive examination of these effects.

Table 1. Reported ratios of intraindividual to interindividual variance in 1-day data

Nutrient	Adult males: Liu data[1]	Adult males: NAS report data[2]	Adult females, NAS report data[2]
	Ratio of variance		
Energy	1.5, 2.0	1.1, 1.0, 0.8	1.4, 0.8, 1.6, 1.6, 1.1
Protein	2.2	1.5, 1.2, 1.4	1.5, 1.3, 2.1, 2.1, 1.4
Carbohydrate	---	1.6, 2.1, 0.6	1.4, 1.2, 1.2
Total fat	2.6	1.2, 1.2, 1.3	1.6, 0.9, 2.1
Saturated fatty acids	2.1, 2.7	1.1, 2.2, 1.4	1.4, 1.7
Polyunsaturated fatty acids	4.4, 5.2	2.8, 3.5, 1.9	4.0, 4.2
Cholesterol	3.0, 4.9	3.4, 5.6, 1.6	4.3, 4.2
Vitamin A	---	(3), 1.6	24.3, 2.5, 7.7, 10.9
Vitamin C	---	3.5, 2.3	2.0, 2.8, 2.3, 2.5
Thiamine	---	2.5, 0.9	4.4, 1.6, 3.3, 3.9
Riboflavin	---	2.4, 0.9	2.2, 1.8, 3.0, 3.3
Niacin equivalent	---	1.6, 2.2	4.0, 2.5
Calcium	1.9	2.2, 1.1	0.9, 1.7, 1.1, 1.2, 1.0
Iron	---	1.7, 1.8	2.5, 1.5, 2.7, 2.5

[1]See chapter 1 of this report.
[2]National Academy of Sciences (1).
[3]All of variance appeared in intraindividual component until logarithmically transformed.

Third, two technical procedures for reducing the true error component of intraindividual variation are proposed: Standardize the collection of data and standardize the coding of data. For the record, Beaton et al. (3) examined the contribution of each of these (not just interviewer effects) in the Lipid Research Clinic's Nutrition System and found, in that highly standardized system, that the contributions were very small, yet intraindividual variation was very large! In unstandardized systems the effect would be greater, and the total variance would be greater. I agree with Dr. Liu's recommendation, but I caution that this is not the basic issue.

Finally, Dr. Liu does not discuss issues of the reliability and precision of the food composition data base and any impact these may have on the estimate of the intake of an individual. Aspects of this have been discussed in a recent NAS report (1) but in relation to a very specific use of food intake data. It is certainly another source of variation that can be either a random variance (that might be factored out in the same way as true day-to-day variation) or a source of bias. It is worth noting that abridged food composition data bases may have an opposite effect: Interindividual variation may be lost to some degree.

Comments on implications and recommendations

Dr. Liu's paper deals almost exclusively with analytical issues arising from the use of proxy estimates of "usual intake." As he points out with examples, two different issues are involved: (a) Observed correlation and regression coefficients may be attenuated, or in some cases exaggerated either way; they will differ from reality; (b) statistical tests for the existence of any relationship will be affected such that power is lost. Thus, there is real potential for an erroneous picture of the nature of the relationship and indeed for a false negative conclusion about the existence of any relationship. These issues, without explicit identification, have plagued the nutrition literature in the past. I see Dr. Liu's paper, at this time, as a plea to avoid this in the future, at least with NHANES III data.

There would seem to be at least three approaches to accomplishing this goal:

1. Increase the reliability of the dietary data through increased number of days of information or selection of a dietary methodology that estimates "usual intake."
2. Attempt to adjust regression and correlation coefficients through application of estimates of the variance ratios.
3. Declare the limitations of the data set and prominently warn users of the errors that must be expected in certain types of analyses.

Dr. Liu recommends the first possibility and suggests that the optimal would be 3 days of data collection. His logic for deriving the optimal number of days would be inadequate for many analyses. Moreover, it is not clear why he is willing to accept telephone interviews in place of face-to-face interviews if the goal is to improve reliability and validity.

Importantly, in his remarks Dr. Liu addressed the matter of selecting an alternate dietary methodology—the food frequency or dietary history. If I may paraphrase him, in selecting one of these methods one is trading error sources, eliminating the problem of day-to-day variation but potentially increasing the problem of imprecision of the estimate. Again, if I understood him correctly, he suggests that a good food frequency method might yield data that are equivalent, for statistical purposes, to three good 1-day recalls or records, but with different types of error sources. Obviously, "bad" methodology of either type would yield bad data! (In a paper published long after this workshop, Beaton (4) demonstrated the apparent loss of a part of the between-person variation in usual intake with food frequency methods. At least part of this loss appeared to be independent of (not correlated with) the true variation in usual intakes. This was additional to any proportional loss of variance that might come from the grouping of foods.) The point, as I understand it, is that switching methodologies does not really get us very far ahead in terms of the statistical issues.

Later in the workshop the question of dietary methodology will be addressed again. However, given the fact that NHANES designers have felt unable to devote more than one-half hour of the survey team's time to the collection of dietary data (see background statements for workshop), I am not optimistic that they are now prepared to invest the time and effort that should be invested to collect reliable or valid data—at least not unless there is a very persuading argument that the benefits to be gained would justify the very real investment that would be needed. This warrants serious attention. If additional investment is out of the question, then our role here is to rule out analyses that cannot be performed. If investment is possible, then our role is to identify the requirements that must be met to provide adequate dietary data. We must have a reading on this if we are to proceed in a meaningful way.

Dr. Liu dismisses out of hand the second possibility, using variance estimates to adjust the observed correlation and regression coefficients. I understand his concerns, but I believe that this is too easy a dismissal. I would argue that, at least in some situations, perhaps those involving only simple correlations or regressions, careful adjustments might be used either to see what the real relationship might look like or even to take into account the intraindividual variation when testing for statistical significance. (In a paper published after this workshop, Rosner and Willett (5) presented an approach to estimating confidence intervals for an adjusted correlation coefficient.) In other applications, in which the distribution of intakes is the focus of interest (as in assessing the prevalence of inadequate intakes in a population or subpopulation), use of variance information to adjust the distribution clearly is in order (1). Again, it must be emphasized that, if this is to be done, the variance estimates must be derived from the data set at hand. That means that there must be a sufficient number of replicate observations distributed in a statistically valid manner to provide reliable estimates of the variances (1). It would not necessarily require three observations of intake to accomplish this purpose.

The third possibility—avoiding the problematic analyses—Dr. Liu does not discuss. Just as it was argued a moment ago that we must know the real constraint to resources for the collection of valid dietary data, we now must know the real purposes of NHANES III. Is it really necessary or cost effective to try to design this survey to accomplish all of the purposes that would involve the type of analyses discussed by Dr. Liu? Must those analyses be addressed in other research data bases collected outside of NHANES?

I would suggest that a major role of the present workshop must be to begin building the matrix discussed earlier, to begin providing the basis for examining questions that might be asked in relation to methodologic, logistic, and financial demands. Only then can the planners of the survey begin making their cost-benefit considerations.

Clearly we face some tradeoffs between potential for use of the data and predictable cost of acquiring the data. The workshop can lay out a framework for examining these tradeoffs, but it cannot make the final judgment.

Toward a matrix of methodology versus analysis potential

Other papers provided as background for the workshop provide a listing of potential applications of NHANES data. One paper (appendix VI) states the purposes of NHANES III as:

1. To monitor changes in health and nutritional status over time.
2. To assess the health and nutritional status of the U.S. population and specific subgroups.
3. To provide information on the interrelationship of health and nutrition variables within population subgroups
4. To measure the prevalence of disease and conditions.
5. To measure met and unmet care needs related to target conditions.

Of these purposes, only purpose 3 carries any mandatory requirement for dietary data. Therefore, at face value it may be fair to state that the purpose of this workshop is to consider the dietary methodology required to permit the examination of the interrelationship of health and nutrition variables within population subgroups, given the following:

- The other purposes of NHANES III cannot be compromised.
- There is another national survey that provides information on dietary intake (but not health conditions).
- There is another set of recommendations that strongly urges that there be sufficient design and methodologic comparability to permit linkage of the two large U.S. data bases—NHANES and the Nationwide Food Consumption Survey (NFCS).

If we then turn to other papers submitted as background to the workshop, we find that there appears to be a strong focus on the ability to link diet (or nutrient intake) to certain specific conditions, notably:

- Cancer.
- Cardiovascular disease.
- Osteoporosis.
- "Energy balance" and obesity.

I think that we will find that linking diet to energy balance has been thoroughly, and properly, dismissed as inappropriate for NHANES (chapter 4). I believe that we will find also that linking diet to osteoporosis has been challenged in terms of any examination of current relationships between intake and bone mass; apparently, a longitudinal study in which serial measures of change within individuals can be measured is recommended (chapter 5). In the case of cancer it appears that what may be wanted is the opportunity to examine intakes today in relation to health of the same individual many years from now—a reliable baseline descriptor of the individual's current intake of either foods or food constituents (it could be either or both). (See chapter 2.) The same could hold for cardiovascular disease; however, here we also face interest in examining current associations among food-nutrient intake, serum lipids, blood pressure, and perhaps such individual traits as smoking, as well as age and sex (chapter 3). This scope of interest is exemplified as a series of questions presented by Dr. Patricia Elmer (chapter 3).

In a separate report, another committee has recommended that the NHANES data base, perhaps combined with the NFCS data base, be used in a joint assessment of nutritional adequacy and nutritional status—assessments of prevalence of inadequate intakes and of prevalence of biochemical and physiological conditions in subpopulations (1). For this purpose the main role of NHANES dietary data might be to validate the agreement of intake between surveys for selected population groups. The critical element would be the ability to link population subgroups through sampling design and subject identifiers and to ensure comparability of at least part of the dietary methodology and of food composition data bases (6). This appears to be a recommendation that will find favor with the Office of Management and Budget and will become a part of a nutrition-monitoring use of the NHANES data base.

It has been urged that national data bases, including NHANES, be used in an epidemiologic mode to attempt to validate estimates of nutrient requirements in the derivation of future Recommended Dietary Allowances (RDA's) (7). This could place special demands on the food intake methodology of NHANES or might require only the ability to link NHANES and NFCS as above.

Dr. Woteki (appendix VIII) gave generalized examples of uses of dietary data:

1. What are the food consumption patterns and nutrient intakes of subpopulations of the United States by such characteristics as age, sex, race, income, occupation, and education?
2. How do nutrient intakes and food consumption patterns of persons differ by level of education?
3. What are the regional differences in consumption of certain food groups?
4. How do nutrient intakes compare with the RDA's and other dietary guidelines?
5. What dietary patterns and nutrient intakes are associated with differing levels of health or health risk?
6. What are the relationships among dietary intake, biochemical indicators, and health status for persons who smoke, use vitamin-mineral supplements, or use oral contraceptives?
7. Do diets of subpopulations with high serum cholesterol levels differ from those with lower levels?
8. What changes in obesity, diet, and activity patterns will take place in the next 10 years?
9. Will diet help explain the continuing decline in serum cholesterol values among men and women?

In her paper, Dr. Woteki describes historical uses of past cycles of NHANES and other surveys as well as predicts possible future uses.

I suggest that many uses of NHANES have been inappropriate in that there has been a more suitable and underutilized data set available: The NFCS data set. It has not always been readily or promptly accessible, but there is at least hope that this will change for the better (1). This is important. A committee charged to examine the two surveys concluded that there were a number of reasons why they could not be combined and why the full scope of each was needed (6). It was seen as critically important that steps be taken to ensure that the data bases could be linked, at least by characteristics of the individuals, so that some questions could be addressed by using the two data bases in concert. It was also seen as important that users be directed to the appropriate data base rather than insisting that either or both be modified to meet the needs of all users (6).

If we then ask which of Dr. Woteki's questions demand the use of an NHANES-type data base, I suggest that we find questions 5–9. If we envisage the ability to link NHANES and NFCS by population characteristics (that is, link subgroups), then only questions 5 and 6 place a specific demand on the dietary methodology of NHANES.

My point, by this time, should be obvious. Any major change in dietary methodology for NHANES III is going to be very expensive. We must be sure what questions demand such a change, then ask what change would be required to address these questions, then finally ask whether this should be done within the framework of NHANES or in some other specially designed survey-research study. By my process of elimination, I am left with the following area of interest and questions that would place special demands on NHANES III and cannot be effectively addressed either by NFCS or by a linking of NFCS and NHANES, including simple dietary methodology.

Applications requiring use of an NHANES-type data base

- To provide information on the interrelationship of health and nutrition variables within population subgroups.
- To provide the opportunity to examine intakes today in relation to the health of the same individual many years from now, a reliable baseline descriptor of the individual's current intake of either foods or food constituents. The retrospective or prospective studies of diet and chronic disease require NHANES III only if the studies require that health data be available at the same time as initial dietary data or if it is impossible to retain subject identifiers in NFCS.
- To be used in an epidemiologic mode to attempt to validate estimates of nutrient requirements in the derivation of future RDA's.
- To determine what dietary patterns and nutrient intakes are associated with differing levels of health or health risk.
- To determine the relationships among dietary intake, biochemical indicators, and health status for persons who smoke, use vitamin-mineral supplements, or use oral contraceptives.

With the exception of baseline data for prospective-retrospective studies, we return to several variants on a similar theme, all concerning the relationships between diet and a health condition or biochemical-physiologic measure assessed only in NHANES. However, if we look at these carefully, not all imply the same demand upon the dietary methodology.

Some could be addressed in simple correlation or regression studies within categorical groups. For these, the possibility of adjusting observed relationships through application of estimates of the variance ratio remains. If so, a 1-day intake data base with appropriate sampling for replicate intakes remains a possibility.

Some of these could be addressed by classifying individuals by the dependent variable (assuming it is more stable) and then asking about differences in intake—for example, in answering the question, "What dietary and nutrient intakes are associated with differing levels of health or health risk?" Here, if sample size is adequate, the 1-day intake might suffice.

Others imply multivariate analyses of one type or another—for example, to ask about the multiplicity of factors that may influence a variable such as serum cholesterol. These then are really the questions that drive methodology issues. These are the questions that must be assessed in terms of their appropriateness for NHANES III.

Table 2 is a first attempt to begin the categorization of questions and data demands. To be useful it would have to be elaborated.

I realize that all of us would wish to be able to conduct quite elaborate analyses on the data set and hence would prefer to have dietary data that are completely adequate for all of our purposes. However, I have the temerity to suggest that NHANES III is not a research study. Although, at considerable expense, we may be able to implement a major change in dietary intake methodology that would seem to meet our purposes, we must ask whether other limitations of the design of NHANES would still constrain the type of analyses we would wish to perform. If so, there was no point in making the dietary methodology change.

The purpose of these remarks, then, is to try to make us think in a broader perspective of what can be done with other data bases alone or in combination with NHANES III, what can be done with a feasible modification of the NHANES dietary methodology, and what cannot be done within the framework of any national survey but must be reserved to the targeted research study. Only when we address this seriously can we really offer effective advice on the design of NHANES III.

Followup remarks

I now take this opportunity to add some comments and suggestions about implications of some of the workshop discussions. I had sincerely hoped that future NFCS and NHANES studies would have enough in common that the types of multiple survey analyses that I portrayed would be feasible, given competent statistical input. I think that was the sincere hope of the original NAS committee (6). From remarks made at the meeting, both formally and informally, I am pessimistic that this will be very effective, at least in the coming round of surveys. If I am correct, this is very unfortunate but perhaps understandable; it was never seen as an easy job to try to achieve

Table 2. Some research questions and applicable data bases

Sample question	*Particular need*	*Appropriate data base*
Association of intake with education, income, or other traits	Group mean intake; traits of individuals	NFCS; NHANES for some traits
Association of intakes of 2 or more nutrients	Reliable estimates of nutrient intake by individuals	NFCS barely adequate; NHANES with greatly enhanced intake data
Association of pattern of food use with nutrient intake	Reliable estimates of pattern of food use by and nutrient intake of individual	NHANES with greatly enhanced quantitative intake data; NFCS with added food frequency
Prevalence of inadequate or excessive intakes	Distribution of "usual" intakes;[1] categorical variables	NFCS; NHANES less satisfactory
Prevalence of biochemical or physiological indices of inadequacy or excess	Distribution of persisting biochemical levels-physiological markers;[1] categorical variables	NHANES (including replicated measures)
Nutrition monitoring (biochemical and dietary)	As in preceding 2 categories	NHANES and NFCS must have linking categorical variables
Do subjects with ______ have different dietary intakes?	Reliable categorical variables; group mean intakes	NHANES for medical-biochemical categorization; NFCS for household, etc., categorization
Role of diet in explaining variance in biochemical-physiological measure	Reliable estimate of persisting intake[1]	NHANES with greatly enhanced dietary data
Baseline nutrient intake for prospective or retrospective study	Reliable estimate of persisting intake	NFCS (barely adequate); NHANES if initial clinical data needed (enhanced dietary data needed)
	Reliable estimate of pattern of food use	NHANES food frequency data, perhaps enhanced
Association of dietary and biochemical-physiological risk factor in individual	Reliable estimate of persisting intake and of persisting biochemical or physiological measure	NHANES with greatly enhanced dietary data
"Epidemiologic" validation of nutrient requirement estimates	As in nutrition monitoring above—comparability of prevalence estimates	NHANES and NFCS (with linking categorical variables)
	As in association study above	NHANES with greatly enhanced dietary data

[1]Anticipates availability of replicated values and opportunity to adjust observed distributions to eliminate any biologically significant effects of day-to-day variation in variable. See National Academy of Sciences (NAS) report (6) for discussion.

NOTES: Dietary nutrient intake questions cannot be asked if food composition data base does not provide reliable estimates of composition of foods with regard to target nutrient. See NAS report (6) for discussion. NFCS is Nationwide Food Consumption Survey. NHANES is National Health and Nutrition Examination Survey.

even the minimal levels of integration and coordination that had been hoped for. However, perhaps more importantly, if this is not achieved in the present round, there are serious implications for both surveys and for those who wish to use the information. I think that everyone stands to lose. Both of the surveys will suffer in terms of interpretations that can be made. Unless NHANES changes significantly from the last round, it will potentially suffer more than NFCS; that is, there is less that can be done with the NHANES II dietary data than can be done with the NFCS past dietary data.

My conclusion, then, is that given my current reading, it is extremely important that NHANES implement at least minimal enhancement of the dietary data, at least a capability to do distributional analyses within NHANES III. The proposal for sufficient replication to do distribution adjustment within the cells of NHANES III data represents such a minimum. With that, I truly believe that there is a very favorable cost-benefit ratio. The cost (dollars and logistics) will go up quite considerably. However, the potential gain in information will go up disproportionately!

I hope that all recognized that in the interchange with Dr. Liu, my minimum enhancement and his were driven by different analytical objectives. If my minimum were adopted, you might be able to adjust some of the simple regressions and correlations, but even that would be statistically undesirable (and would really do little or nothing to increase the statistical power to establish the presence of effects; its main effect would be on the description of the effects). It really does little to address the issues put forward by Dr. Liu in his excellent paper. Dr. Liu is correct. Three days would be better for that purpose (and of course would be better also for the simple distribution adjustment), but I am not convinced that 3 days would be good enough for the other analyses to be worth the additional investment. This, obviously, is something that you and your colleagues will have to address: What power can you gain, and hence what additional questions can you answer, for the additional resource investment (dollars or logistics) that would be required? My bias is toward a different instrument for those types of questions. My concern is that, although I can see and have an intuitive feeling for the obvious gain to be obtained from reducing intraindividual variance through use of one of the several variants on a food frequency instrument, I cannot get any intuitive feeling for the impact of the loss of precision of identification of foods and amounts involved in such an instrument. There is a tradeoff in errors and I cannot assess it. It might be worth contracting for someone to undertake a paper, such as Dr. Liu's, or to undertake "sensitivity analyses" of the type undertaken in the NAS report (1; 8, pp. 137–159) to try to get a handle on the pros and cons of those instruments. You are fully aware of the fact that, although it is difficult to validate the 1-day intake instruments, it is virtually impossible to validate the frequency-history instruments. (How do you obtain an independent estimate of "usual" intake?). Perhaps, then, the only approach is one of examining the potential effect of adding and/or removing the different types of errors rather than attempting to establish "truth."

For the record and based on my empirical work adjusting the NFCS data sets with 3-day data, the adjusted and actual distributions are not terribly different. That is, as many people have shown, the observed variance in population data falls sharply as one adds days from 1 to 3, but after that the reduction in variance diminishes (the basis of Dr. Liu's arguments for 3 days). This does not give reliable estimates for each individual but does raise the interesting point that for distributional analyses, one should consider the cost-benefit ratio of adjusting the distribution during analysis versus using the observed distribution. Either way, the critically important point is that, although you can estimate proportions of the population falling within the defined intervals, you cannot classify the particular individuals falling in these intervals without appreciable error.

I wanted to emphasize another aspect of the NAS report (6). Although that report's approach to adjusting the distribution of intakes was a necessary prerequisite, it was really a minor part of the report and its implications. The major thrust was in relation to the use or nonuse of cutoff points for dietary (or biochemical) data. I think that you and your colleagues have been sensitive to the issues surrounding the cutoffs in biochemical data. I am not sure that you have really solved the issues. I am certain that you have not resolved the issues of interpretation of dietary data. The NAS report was clearly a beginning, not an end. It provides a construct that I think is meaningful. However, it now needs a group to address the question of nutrient requirements or toxicological "requirements" defining in particular average requirements for defining states of nutriture. This was never (in recent years, at least) a goal of the RDA committees; hence, you will not find the estimates needed within the RDA reports. The development of such estimates is of importance to both the U.S. Department of Agriculture and the National Center for Health Statistics if they are to interpret the next round of surveys. My message, then, is that there is merit and mutual benefit if the two agencies can blend forces in an attempt to have some group address this specific goal in the next year or two—without getting it involved in the controversies of meaning and derivation of the "RDA's." These really are two different things.

(Since the workshop, the Food and Agricultural Organization and the World Health Organization have published or initiated three reports dealing with nutrient requirements and/or population nutrient goals. In these reports, the committees have been much more explicit in defining the "what" in "requirement for what," and indeed multiple levels of requirements for different nutritional states have been presented. The committees have been explicit also in addressing the statistical issues of population intakes versus individual intakes and in differentiating between the low-risk individual intake (the conventional RDA) and population mean intakes consistent with a low prevalence of inadequate or excessive nutrient intakes. It is these types of definitions and specifications that are needed in the analysis and interpretation of data sets such as NHANES and NFCS.)

Again, my thanks for inviting me to the meeting; it was an interesting and informative session. I wish you and your colleagues every success with the important task that now lies ahead in designing NHANES III.

References

1. National Academy of Sciences, National Research Council, Subcommittee on Criteria for Dietary Evaluation, Coordinating Committee on Evaluation of Food Consumption Surveys. Nutrient adequacy: Assessment using food consumption surveys. Washington: National Academy Press. 1986.
2. Morgan KJ, Johnson SR, Goungetas B. An analysis of a 12-day data series using persistence measures. Am J Epidemiol 126:326–35. 1987.
3. Beaton GH, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J Clin Nutr 32:2546–59. 1979.
4. Beaton GH. Interpretation of results from diet history studies. In: Kohlmeier L, ed. The diet history method. Proceedings of the second Berlin meeting on Nutritional Epidemiology. London: Smith-Gordon. 1991.
5. Rosner B, Willett WC. Interval estimates for the correlation coefficient corrected for within person variation: Implications for study design and hypothesis testing. Am J Epidemiol 127:377–86. 1988.
6. National Academy of Sciences, National Research Council, Coordinating Committee on Evaluation of Food Consumption Surveys. National survey data on food consumption: Uses and recommendations. Washington: National Academy Press. 1984.
7. National Academy of Sciences. Unpublished report of a workshop to plan a new Recommended Dietary Allowance. Washington: National Academy of Sciences. 1985.
8. Beaton GH. New approaches to the nutritional assessment of population dietary data. In: Proceedings of the Tenth National Nutrient Data Bank Conference. U.S. Department of Commerce. 1986.

Chapter 2 Dietary assessment issues related to cancer for NHANES III

by Gladys Block, Ph.D., National Cancer Institute

It has been estimated by Doll and Peto (1) that 35 percent of human cancer may be associated with diet, although the range of reasonable estimates is very wide, from 10 to 70 percent in their estimation. In addition to suggestive international data on correlations between diet and disease, numerous epidemiologic studies have found associations between dietary factors and specific cancers. The National Health and Nutrition Examination Survey (NHANES), of the National Center for Health Statistics (NCHS), offers the unique opportunity to learn what people are eating, essential to the design of effective health promotion programs; to study changes over time and evaluate the success of large-scale interventions designed to alter dietary habits; and to evaluate the relationship of dietary factors to physiological-biochemical factors and health outcomes, including cancer.

The first two of these functions, assessing dietary consumption and changes in consumption over time, can be performed as well by the Food and Drug Administration and the U.S. Department of Agriculture (USDA), chiefly through the USDA's Nationwide Food Consumption Survey (NFCS). The existence of that sister survey is important to bear in mind, for it has important consequences for the methodologic choices that must be made in the design of the third National Health and Nutrition Examination Survey (NHANES III). NFCS, as its name implies, is designed to assess food consumption; NHANES, on the other hand, is the only large national survey designed and mandated to assess not only dietary intake but the relationship between diet and health. To quote from the NCHS document on the plan and operation of the second NHANES (NHANES II): "The essential differentiating characteristic of the health examination surveys is their primary concern with those kinds of health-related data obtained only (or at least optimally) from direct medical examinations . . . . The NHANES surveys in their present form were designed to permit relating nutritional variables to health measures" (2). This has important methodologic implications for the nature of the dietary assessment methodology required to achieve this goal.

Specific dietary constituents and cancer

There is laboratory or epidemiologic support for, and current research interest in, the relationship between cancer and a number of dietary constituents, including the following: Total calories; total fat; percent of calories derived from fat; specific fat fractions such as polyunsaturated fatty acids; cholesterol; dietary fiber and fiber fractions; vitamins A, C, and E; carotenoids; folacin; trace minerals including selenium and zinc; alcohol; and others. However, it is important to recognize that a few years ago the list would have been different; a few years in the future the list will almost certainly be different. Consequently, it is unwise to focus our attention on a limited list of a few nutrients. Rather, to be maximally useful even to the state of knowledge that will exist by the time the survey itself is completed, the dietary assessment must be designed to capture, to the best of our ability, the whole diet rather than a limited number of nutrients. Only by capturing the whole diet can new hypotheses be generated and tested, and existing hypotheses tested adequately, and only by capturing the whole diet can we begin to examine dietary interactions, a potentially important area about which our understanding is minimal at best.

Dietary assessment methods

These have been reviewed extensively elsewhere (3,4) and will only be summarized here.

24-hour recalls

In this method, the detailed dietary intake over the past 24 hours is reported. It should be used with three-dimensional models to enhance portion size accuracy, as research indicates that individuals of all ages are poor estimators of their portion sizes, and differing portion size assumptions also constitute an important source of discrepancies among coders and data bases. In addition, interviewer probes should be detailed enough to preclude "not further specified" codes. The latter are an important source of imprecision in group estimates, especially if the nutrient values for the "not further specified" code are assigned by the so-called "conservative" approach, which uses the most extreme value in the data base (the fattiest cut of meat, for example), regardless of its frequency of consumption.

The 24-hour recall has been used in previous NFCS and NHANES surveys for children and infants (by means of surrogate respondents) as well as adults. It provides valid data on the mean intake of groups or the intake of an individual in the prior 24 hours. It does not provide valid estimates of the usual diet of individuals over a longer period of time (5–7).

Diet records

In this method respondents, in principle at least, record their diet as they are consuming it, thus minimizing errors of memory. Conscientiously applied, it can capture components of diet that might be overlooked by a recall and affords the opportunity to obtain more accurate portion size estimation through measuring or weighing. It clearly cannot be applied to children, whose diet must again be recorded by a surrogate. It requires respondents who are reasonably literate and comfortable with the idea of careful recordkeeping and who are sufficiently committed to the goals of the survey to perform the recordkeeping conscientiously. Like 24-hour recalls, it can provide valid data on groups; its validity for the individual's usual intake depends on the length of the record, discussed below.

Multiple days of recalls or records

This approach is an attempt to obtain dietary data more representative of the individual's usual intake than is a single day's intake. This issue is discussed further below. Multiple-day recalls should be obtained as a series of 24-hour recalls rather than as a single 3-day recall, for example, as there is a serious decrement in memory beyond the previous 24 hours (8). Furthermore, some data indicate that the accuracy and completeness of multiple-day records suffers after the first 2 of 3 days, at least in some population groups (9).

Frequency questionnaires

In this approach the respondent is asked to indicate his or her usual frequency of consumption of each of a list of foods, or sometimes food groups. This may be quantified in a variety of ways in which a weighting factor such as nutrient quantity in a "standard portion" is multiplied by the frequency of consumption to obtain a nutrient score or estimate of the quantitative nutrient content of the diet. The quantitation can also be designed so that the "standard portions" are age- and sex-specific, thereby improving somewhat the precision of the nutrient estimate.

The food list has, in the past, often been targeted toward the assessment of a single nutrient, such as vitamin A, or a few nutrients, and then would include only foods designed to capture that nutrient. That is not an inherent feature of frequency questionnaires, however, and food lists can be designed to capture all of the important nutrients in an individual's diet. To do so, the food list must reflect the population's major sources of those nutrients, and the food items must be kept reasonably distinct. A frequency questionnaire containing an extensive food list designed to capture a broad range of nutrients in an individual's diet might then differ from a list-based diet history solely in that the list-based diet history is designed to obtain some measure of the respondent's usual portion size of each food, and the frequency questionnaire is not. Validations are discussed below.

Diet histories

This term has been applied to a wide variety of approaches. What I take to be the defining characteristics of a diet history are that it is designed to elicit quantitative information about the individual respondent's usual dietary intake over an extended period of time, with an attempt at quantitation of the individual's usual portion sizes of the foods in his or her diet. The quantitation may require the respondent to report the usual portion size of each food in household measures, or the respondent may select a three-dimensional model or a photograph that most closely resembles the usual portion. As it is usually defined, a diet history involves probing for specific food types, portion sizes, and seasonal differences, requiring a 1- to 1½-hour interview with a trained nutritionist. However, if it is list based and structured, a shorter time and nonnutritionist interviewers may be used.

Although Burke's original method (10) was primarily menu based, with lists and recalls used for cross-checking, many current approaches to diet history interviews are primarily list based; that is, the individual is prompted with a list of specific foods, and then the frequency of consumption and portion size are obtained for those foods reported as consumed. Menu-based approaches may be heavily influenced by the individual's diet of the past few days (11) and may be distorted by people's inclination to report three "normal" meals per day, when in fact modern diets are less orderly. List-based approaches, on the other hand, are dependent on the adequacy of the list of foods. Validity is discussed under "Critique of methods for NHANES III," below.

Appropriateness of methods

For children, all of the methods encounter the same problems: The questionable ability of the child to respond accurately, and the questionable ability of a surrogate to report what the child may have eaten outside the home. If a diet history or quantified frequency approach were used in NHANES III, preliminary work would be required to determine the food list appropriate for capturing the diet of children. Extensive work on this subject is under way at the University of Texas for the Women, Infants, and Children (WIC) program (12). Similarly, an optimal diet history or quantified frequency questionnaire would require information on children's usual portion sizes of each of the foods; this could be obtained from the NHANES II data, which used three-dimensional models.

For some nutrients, all dietary methods are probably inappropriate. The selenium content of grain, for example, depends on the soil in which it was grown, and soils in the United States differ markedly in their selenium content. Such nutrients are probably better assessed in biological samples.

The appropriateness of methods also varies for different study designs and goals, as follows.

1. *Determining the adequacy of the dietary intake of the population and its subgroups*—A major function of large national surveys is to determine the population's intake of macronutrients and micronutrients and to assess the adequacy of intake of subgroups defined by age, race, poverty, or other demographic characteristics. For this purpose, information is required that is accurate at a group level, and it must be precise. Whether a group's mean intake of vitamin C is 50 milligrams or 60 milligrams has rather important implications for national health and food policies; it is probably not sufficient simply to know that one group is higher than another in dietary intake. The most efficient and valid method for obtaining accurate group data is a carefully performed 24-hour recall, as done in previous NHANES.

However, it is important to note that (a) this is not the only, and perhaps not the primary, role of NHANES, as suggested in the quotation above; (b) it can be performed adequately by the other national survey, USDA's NFCS, for which this clearly is the primary mandated purpose.

2. *Determining temporal trends and comparability with prior NHANES*—To be fully comparable with prior NHANES data, NHANES III and future surveys would have to use identical methodology—24-hour recalls. However, even if 24-hour recalls were used, the data would still not be fully comparable, as a decision has been made to use the USDA food data base rather than the NHANES data base. The latter included considerable manufacturer data and food codes for mixed dishes that either do not exist or are calculated differently in the USDA data base. Thus, even an NHANES III that used one or more 24-hour recalls would be only partially, not fully, comparable with prior surveys.

If a frequency questionnaire or diet history were used, there would be uncertainty as to whether apparent changes in group mean intake were due to real dietary changes or to methodologic differences. It is possible, however, to construct these methods so as to provide group mean estimates that are quite closely comparable to those of NHANES II (13). This is done by making use of the food-specific portion sizes reported by NHANES II respondents using three-dimensional models. Table 1 shows group mean caloric data for two sex and five age categories, from the NHANES II survey, and from the National Cancer Institute's list-based diet history questionnaire (13). None of the list-based diet history group mean estimates differs statistically significantly from NHANES II estimates. Moreover, the ranking of the age-sex groups and the large range in intake is well preserved. Similar data are seen for the other nutrient measures. Thus, it may be possible to evaluate changes in group means even if the methodology is changed, provided it is developed appropriately.

It is important to evaluate the costs, both monetary and opportunity costs, of attempting to preserve strict

Table 1. Mean calorie values from NHANES II and list-based diet history (LBDH) (including both self- and interviewer-administered values), by sex and age

Sex and age	Number	NHANES II Mean	NHANES II Standard error	LBDH Mean	LBDH Standard error
Male					
25–34 years	34	2,746	52	2,624	141
35–44 years	107	2,434	47	2,328	66
45–54 years	121	2,371	48	2,237	69
55–64 years	146	2,076	26	2,068	54
65–74 years	125	1,818	24	1,851	59
Female					
25–34 years	98	1,633	25	1,698	49
35–44 years	42	1,584	29	1,620	79
45–54 years	36	1,443	27	1,608	93
55–64 years	64	1,402	19	1,488	55
65–74 years	34	1,301	18	1,408	84

[1]NHANES (National Health and Nutrition Examination Survey) II values were calculated using SESUDAAN to adjust for sample design.

SOURCE: Block et al. (13).

comparability by retaining the 24-hour recall and to consider other sources of time-trend data. This is exactly the kind of information that is obtainable from the USDA's NFCS. Its design and mission allow for the collection of group mean data comparable from survey to survey. On the other hand, the use of 24-hour or few-day data in NHANES III would vastly diminish the value of all of the examination and laboratory data and subsequent health followup, because it would not be possible to relate the individual's health or examination and laboratory measurements to dietary intake with anything like the power that individually valid data would provide. It is exactly considerations such as these that led the NCHS and National Instutotes of Health investigators designing the NHANES I Epidemiologic Followup Study to employ a food frequency approach, so as to have data on an individual's long-term usual intake, which is more relevant to health outcomes. Thus, although precise intake and time-trend data for groups may be obtainable only with 24-hour or 3-day data, the existence of the other large national surveys makes it possible to consider other methods for NHANES III to better serve its unique health assessment roles.

3. *Using cross-sectional studies*—An important role of NHANES is to provide information about the joint distribution of demographic, dietary, and physiologic characteristics in the population and to provide information for analyses of the relationship between dietary and physiologic factors. Some studies of this sort can be conducted at the group level and thus could be supported by 24-hour or few-day data. It is possible, for example, to obtain useful information from a trend analysis in which the mean dietary intakes at successive levels of income or age are examined. However, as Lilienfeld and Lilienfeld have said, "In studying group characteristics, the results of such studies provide clues to etiological hypotheses" but "may

suffer from an 'ecologic fallacy'.... Associations established by studying characteristics of population groups have a greater chance of being indirect than those established by studying characteristics in individuals. In other words, it is preferable to have an association based on data obtained at the level of the individual and not at the level of a population or group" (14).

An analysis of dietary and serum cholesterol is a case in point. Studies of the NHANES I and II data have found little or no relationship between dietary and serum cholesterol (15,16). However, such a relationship could hardly be detectable in those data because 1-day dietary data are largely irrelevant to an individual's longer term intake. Thus, to be maximally useful for cross-sectional studies, dietary data for NHANES III should reflect the individual's usual dietary intake; 24-hour or few-day diet information does not achieve this purpose.

Analytic epidemiology—cohort and case-control studies

In order to understand and evaluate the relationship between a dietary factor and a disease, cohort studies (prospective) and case-control studies (retrospective) are performed. In the former, individuals are characterized with respect to an exposure, and their subsequent disease outcome is examined. In the latter, individuals are defined with respect to the presence or absence of a disease and then characterized with respect to their exposure to a factor at some time in the past. For both types of studies, it is important to note that individuals must be characterized or classified with respect to their exposure to a factor so that they can be placed in, for example, quartiles of the distribution; that is, the individual's true usual level of intake of a dietary factor must be known. Group mean values are of little value.

Consequently, if the dietary assessment method used in NHANES III is capable of assessing the individual's usual diet, this survey will lend itself to the evaluation of hypotheses regarding the possible relationship between dietary factors and disease outcomes or physiologic states. Dietary methods that cannot be used to at least place individuals into the correct quartile or so of the distribution of intake of a nutrient will not permit studies of this sort with any statistical power.

Prospective studies can be conducted within NHANES III if the appropriate identifier data are collected; individuals would be categorized with respect to their usual intake at baseline and their subsequent health outcome tracked. Retrospective studies would be conducted by waiting until health outcomes have been tracked at some time in the future, identifying cases of a disease and controls, and then looking back at and classifying them with respect to their usual intake at baseline. Thus, identical data are needed for the two types of studies.

It is likely that in relation to the initiation or even promotion of cancer, the diet at some time in the past is what is relevant. Although some methods can provide moderately good estimates of past diet (17–19), such methods would be inappropriate in NHANES III. It would be inefficient to attempt to collect information on past diet from all respondents, when only a small subset of cases and controls will be used. Rather, the current usual diet collected by NHANES III will become the "past" diet when cases and controls are identified in the future.

Critique of methods for NHANES III

24-hour recalls

In my view it would be a grave mistake for NHANES III to use only a 24-hour recall, either with or without a subsample in which two recalls were collected. Despite the benefits of 24-hour recalls with respect to group data and comparability, the opportunity costs are just too great to justify their use—costs such as the inability to make inferences about the relationship of diet to health or to physiologic factors such as serum cholesterol. All of the health measures of interest, all of the sociobehavioral factors, and all disease outcomes are maximally interpretable only in relation to the usual diet of an individual. There is ample evidence that, although 24-hour recalls can adequately reflect the mean values of groups (and even that is questioned by James, Bingham, and Cole (20)), they are unrepresentative of an individual's usual intake (5–7). This failure is not relieved by limiting analysis to the proportion of various nutrients in relation to reported caloric intake. Indeed, in Beaton et al. (5) data, the ratio of intraindividual to interindividual variability actually increased when nutrients were expressed as a percent of energy. They indicated that "the elimination of the effect of energy intake by calculation of the ratio of nutrient:energy would be expected to reduce variability. Nevertheless, the relatively high intraindividual variability suggests wide variation, day by day, in the composition of self-selected diet." This position is also supported by James, Bingham, and Cole (20), who concluded that "an individual's variable intake of protein and energy is sufficient from day to day to limit the possibility of using nutrient density rather than absolute intakes as discriminators between individuals."

Multiple days of recalls or records

If 1 day provides an inadequate characterization of an individual's usual intake, the next logical suggestion is to increase the number of days. Thus, NFCS and numerous other studies use 3 consecutive days of diet, either by record or a combination of record and recall, in an attempt to achieve both quantitative precision and greater representativeness of usual diet. How many days, then, are required to obtain adequate representativeness of an individual's usual diet?

This question has been studied intensively in the last decade, but it goes back at least as far as Chalmers et al.

(21) in 1952. Chalmers et al. came to the same conclusions as more recent investigators such as Beaton et al. (5,6), Liu et al. (22), and Sempos et al. (23); it remains to be seen whether the later conclusions will be heeded any more than the earlier ones. Their data all indicate that a small number of days of diet information does not represent an individual's longer term usual intake with any precision. Indeed, an individual's diet is sufficiently variable from day to day that intraindividual variability is greater than interindividual variability for virtually all nutrients studied. The results are extremely consistent across all of these investigators and their different study populations.

Beaton et al. (5,6) found that to be within ± 20 percent of the true usual intake (a rather generous criterion, representing, for example, between 1,600 and 2,400 for a true usual intake of 2,000 calories) would require 7 days for males and 10 days for females. For protein, it would be 13 days for males and 10 days for females. For percent of calories from fat, the best case in Beaton's data, it would require 3 days for males but 5 days for females. For most micronutrients the situation is even worse. If one wanted precision to within ± 10 percent, a great many more days would be required.

Liu et al. (22) examined the ability to classify individuals correctly into quintiles of the distribution. To have less than a 5-percent chance of grossly misclassifying someone from the first to the fifth quintile of cholesterol intake, again a rather generous criterion, would require 7 days of diet information. To avoid misclassifying someone from the first to the third quintile or vice versa would require 11 days.

Sempos et al. (23) used a somewhat different criterion. Suppose one wanted to ensure that an observed correlation between a dietary factor and a physiologic factor such as serum cholesterol was at least as much as 0.9 of the real correlation. This would take 7 days for calories, 9 days for protein, 6 days for calcium, and at least 15 days for micronutrients such as the B vitamins.

Other investigators have found similar results. Balogh, Kahn, and Medalie (24) found that to be within ± 20 percent for 90 percent of the population would take 9 days for calories and 10 days for cholesterol. James, Bingham, and Cole (20), in England, found that to correctly classify 80 percent of men into the correct third of the distribution would require 5 days for calories, 9 days for fat, and 18 days for percent of calories from fat. Finally, to give the 1952 paper of Chalmers et al. its due (21), they calculated that to have a 95-percent confidence interval equal to ± 15 percent of the then Recommended Dietary Allowance for calories, one would need 14 days for men and 11 days for women.

As Beaton has asked (25), "What do we think we're measuring?" Three days look good because they give us reasonable means (which even 1 day can do) and because we have not looked at the intraindividual variability of such 3-day estimates or their representativeness for a larger number of days in an individual. Balogh, Kahn, and Medalie (24) found that the more days of data that were collected, the more intraindividual variability was revealed. In data of Willett et al. (26), even two 7-day records were correlated with each other at only about 0.5.

In most of the above data I have quoted the best cases, the macronutrients that one would suppose were reasonably stable from day to day or over a few days. For most of the micronutrients the situation is worse, and investigators have found 10, 15, 20, or more days to be needed. For vitamin A, the situation is totally hopeless.

In addition to the issues of precision and representativeness, multiple-day data, as they would probably have to be collected in NHANES III, would not relieve the serious confounding of season and region caused by logistical necessity. In previous designs, at least, the southern regions were examined in the winter and the northern regions, in warmer seasons. The result is that the Northeast, examined in the warmer months, appears to have higher intakes of vitamins A and C and more consumption of fruit, for example, than does the South, examined in the opposite season; and the Midwest and West, spanning both northern and southern areas, have intermediate values. The result is that few valid inferences can be drawn about regional differences in dietary intake, nor can geographic correlations regarding diet and disease be drawn. Several-day records, collected during the time period of the examination, would only perpetuate this confounding of season and region.

My reading of the above data is that no number of days that is within the range of feasibility for NHANES III is capable of providing yearlong representativeness or nutrient estimates that one can be confident are within ± 20 percent of the true usual intake of an individual.

Frequency questionnaires

Although frequency questionnaires have often been targeted to a few nutrients, they can be designed to obtain a much broader picture of the individual's diet. The questionnaire designed by Dr. Willett at Harvard is an example of such an instrument.

Are such instruments valid? That is, can they either categorize individuals along the distribution of intake and/or provide accurate quantitative estimates? The question should not be given a general answer, because the performance of a questionnaire depends profoundly on the exact food list, as well as the portion size and nutrient content assumptions. Observed correlations may be as much a function of imperfect portion size assumptions made by the investigators, for example, as a function of the respondent's ability to report his or her usual diet. Moreover, the appropriateness of the reference method must be carefully examined. Given the above data on day-to-day variability, it is unlikely that a 7-day record, for example, is an appropriate "truth" measure for usual diet, against which a test instrument is to be judged. (In Beaton's data, for example, only percent of calories from carbohydrate and fat could be obtained from both men and women in 7 days to an accuracy of ± 20 percent.)

Only Dr. Willett's validation study approaches a number of days (four 7-day records) that might adequately represent usual intake for nutrients other than vitamin A. Just as Willett's correlations were poorest for dietary vitamin A (for which the fault might lie with the unrepresentativeness of the "reference" method rather than the questionnaire), so also correlations seen in other studies using shorter reference periods may reflect unrepresentative reference periods rather than unreliable questionnaires.

Surprisingly little research has been carried out on the validity of the nutrient estimates resulting from pure frequency questionnaires, that is, frequency questionnaires that, although they may be quantified, do not attempt to ascertain the respondent's portion sizes. Virtually all of the abbreviated questionnaires usually cited in this connection actually included some quantitation of the individual's usual portion sizes. Jain et al. (27) and Balogh, Kahn, and Medalie (24), for example, used pictures of different portion sizes. Browe et al. (28) had the subject "compare his usual portion with (a stated standard portion) and make appropriate changes on the questionnaire." Reshef and Epstein (29) also state that they collected "the size of the usual portion." Several authors have examined the reliability of the frequency reports themselves (for example, number of times consumed carrots) on repeat administrations or by spouse pairs. It is difficult, however, to evaluate the meaning of such agreement for the validity of actual nutrient estimates.

Perhaps the only validation of a pure frequency instrument (that is, one that does not entail variable portion sizes) aimed at assessing a wide range of nutrients and using a reasonable reference measure is that of Willett et al. (26). They administered a 61-item questionnaire and compared it with the mean nutrient intake from four 7-day diet records obtained over the course of a year.

The quantitative nutrient estimate was obtained by multiplying the reported frequency by the nutrient amount obtained in a portion, which was stated on the questionnaire (such as "cream cheese, 1 oz."). In principle, respondents could take their own portion size into account to some extent by modifying their reported frequency. Thus, if respondents felt that they usually had 2 ounces of cream cheese instead of 1 ounce, they could change their response category from "2–3 per week" to "4–6 per week." As it seems unlikely that very many respondents actually would do this carefully, I have included the instrument under frequency questionnaires rather than diet histories.

Correlations between the food frequency questionnaire and diet record nutrient estimates for the prior year were approximately 0.45 without adjustment for reported caloric intake and about 0.53 after adjustment for calories. (It should be noted that Willett's study was done in a self-administered format. It is likely that correlations would have been somewhat higher if it had been administered by an interviewer.)

Agreement in this range is greater than chance and indicates a fair ability to categorize individuals at the extremes. For example, of individuals in the lowest quintile by diet record, 40–50 percent were in the lowest quintile by the food frequency questionnaire. Correlations in this range are also comparable to the degree of precision that could be obtained by a few days of diet records. However, they also involve a considerable degree of misclassification. The consequence of this misclassification is that correlations in this range require a fivefold increase in sample size to detect a statistically significant risk trend (29), and the observed risk ratio would nevertheless still be an underestimate of the true risk ratio.

It may be that for some nutrients a pure frequency questionnaire would be adequate to rank individuals reasonably accurately along the distribution of intake, all that is really needed for analyses of diet-disease relationships. Samet, Humble, and Skipper (30) make a reasonable case that vitamin A is such a nutrient. However, it seems clear that this is not the case for most nutrients and that some measure of individual portion size is needed.

Diet histories

Some forms of diet histories require a 1- to 1½-hour interview by a nutritionist. Time constraints make it clear that this form of diet history is not feasible for NHANES III. Several investigators, however, have developed brief diet history questionnaires designed to gather data (a) representative of an individual's usual diet and (b) quantified based on information about the respondent's usual portion sizes. If such approaches are structured and list based, they may provide reasonably accurate quantitation about an individual's usual diet in a shorter interview time.

The validity of such methods as accurate estimators of an individual's usual intake is not easy to evaluate, as "usual" encompasses a longer period of time than most of the methods against which it has been tested. Burke et al. (31) and Reed and Burke (32) evaluated the method against clinical criteria and found excellent agreement. When compared with 7-day records, some but not all investigators have found that the history method gives higher mean values, and correlations have generally ranged from 0.4 to 0.8. Some validation studies are summarized in table 2.

Balogh et al. (11) found correlations ranging from 0.6 (for total carbohydrate) to 0.8 (for fats) between the mean of 8 or more days over a year and a diet history based on 69 food items. Jain et al. (27) validated a different 69-item questionnaire, quantified by the respondent's choice among sets of photographs of different portion sizes, against a detailed 1-hour frequency interview. They reported correlations of 0.6 to 0.7 for nine macronutrients, vitamin C, and fiber. The interview itself, however, when validated against multiple diet records among 16 men, achieved correlations of only 0.24 to 0.61, averaging 0.45 (33). Hankin, Reynolds, and Margen (34) found similar results for a short diet history questionnaire, targeted toward dietary fat, in which respondents selected photographs of small, medium, and large portions.

Table 2. Validation studies of list-based diet histories (LBDH)

Validation study	Pearson correlation
Balogh et al. (11), LBDH vs. 7-day weighed records, macronutrients.	0.7–0.9
Balogh, Kahn, and Medalie (24), LBDH vs. 8+ 24-hour recalls, macronutrients	0.7–0.8
Jain et al. (19,27), 69-item LBDH vs. interview, macronutrients, micronutrients	0.6–0.7
Hankin, Reynolds, and Margen (34), LBDH vs. interview, fats	0.6–0.7
Block, et al. (37,39,40), Cummings et al. (38), macronutrients, micronutrients	0.5–0.7

Other list-based diet histories have been developed by other investigators. Byers et al. (35), for example, describe a 128-item questionnaire. Although they usually administer it as an extensive interview, they have developed shorter versions targeted toward specific nutrients. Similar extensive instruments have been used by others (36), and some might lend themselves to modification so as to fit within the time limit of NHANES III.

We at the National Cancer Institute (NCI) have also developed a list-based diet history (called "Self-Administered Diet History" in the methods paper (13) but capable of being interviewer administered as well). By using a data-based approach, selecting foods on the basis of their population contribution to calories and 17 other major nutrients, and keeping the food items reasonably distinct, the instrument is designed to capture all of the major nutrients in the diet with the most efficient food list. The list contains foods representing at least 90 percent of U.S. dietary intake for energy and each of the 17 nutrients in the NHANES II data base. In addition, portion sizes were developed on the basis of age- and sex-specific data from NHANES II, thus making use of large-population data based on three-dimensional models. The effectiveness of this approach can be seen by the good agreement in table 1 and the ability to distinguish groups by age and sex. Individuals are asked to quantify their usual portion of each food by indicating whether it is small, medium, or large with respect to a stated medium portion. However, other approaches could be envisioned, such as having the individual select a photograph or three-dimensional model for each food, as is done in the list-based diet histories of Hankin, Reynolds, and Margen (34), Jain et al. (27), and others.

The dietary intake of special demographic groups or individuals can be captured more precisely by the inclusion of an open-ended section in which foods frequently eaten but not on the main list can be reported. The questionnaire is administered in an average of 25–35 minutes. Furthermore, if the list is developed on the basis of major nutrient contributors, much shorter lists can yield correlations of 0.90 to 0.98 with the longer instrument and correlations with reference data similar to those achieved by the longer instrument, thus permitting assessment of a wide range of nutrients in a brief questionnaire (37).

Validations of this NCI questionnaire have revealed correlations of approximately 0.6 with various reference methods, as well as considerable accuracy in the estimation of absolute nutrient levels (37–39). Excellent correlations with serum carotenoids have also been found (40).

Thus, the work of Hankin, Jain, Balogh, and others, as well as our work at NCI, all suggest that correlations in the range of 0.6 to 0.8 can be achieved with list-based diet histories, quantified either with photographs or with a "small-medium-large" approach.

Recommendation

The above discussion leads me to the following conclusions. Because NHANES was designed "to permit relating nutritional variables to health measures" (2), NHANES III needs dietary data representative of an individual's usual intake over an extended period of time in order to examine diet-disease relationships. Data based on 24-hour recalls are thus totally inadequate for some of the major goals of NHANES. Multiple days of recall or records, up to at least 14 days or so, provide the illusion of precision with representativeness, but not the reality. Pure frequency questionnaires, in which no individual portion size information is collected, probably provide less precision in nutrient quantitation than do diet histories and thus provide less ability to categorize individuals along the distribution of intake, at least for most nutrients.

I believe the data are quite clear that list-based diet histories, developed and quantified appropriately, can provide the individual representativeness needed for the health goals of NHANES III. For some nutrients, at least, they can also yield nutrient estimates that accord well with known diet or more precise group methods (13,17,38,39, and table 1).

Other important societal goals, such as accurate group mean estimates and measures of changes in intake over time, can be provided by the 3-day recall or record format used in the USDA's Nationwide Food Consumption Survey. In this way the large national surveys would complement rather than duplicate each other, and the major goals of both could be fully achieved.

I conclude therefore that NHANES III should use the method that achieves the greatest degree of representativeness of the individual respondent's usual diet so as to maximize the usefulness of the data unique to NHANES—the physiologic and health data. If 7 to 14 (or more) days of diet recalls could be secured reliably from all respondents in two different seasons, this might be the method of choice. Because this is probably not logistically or economically feasible, I believe there is clear evidence that a list-based diet history, with some quantitation of individual portions, can achieve a degree of representativeness that will permit meaningful analytic studies of the relationship among diet, physiologic measures, and health.

References

1. Doll R, Peto R. The causes of cancer: Quantitative estimates of avoidable risks of cancer in the United States today. New York: Oxford University Press, Inc. 1981.
2. McDowell A, Engel A, Massey JT, Maurer K. Plan and operation of the second National Health and Nutrition Examination Survey, 1976–80. National Center for Health Statistics. Vital Health Stat 1(15). 1981.
3. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 115:492–505. 1982.
4. Burk MC, Pao EM. Methodology for large-scale surveys of household and individual diets. Home Economics Research Report no. 40. Washington: U.S Department of Agriculture. 1976.
5. Beaton GH, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J Clin Nutr 32:2456–9. 1979.
6. Beaton GH, Milner J, McGuire V, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation – Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 37:986–95. 1983.
7. Todd KS, Hudes M, Calloway DH. Food intake measurement: Problems and approaches. Am J Clin Nutr 37:139–46. 1983.
8. Adelson S. Some problems in collecting dietary data from individuals. J Am Diet Assoc 36:453–61. 1960.
9. Gersovitz M, Madden JP, Smiciklas-Wright H. Validity of the 24-hour dietary recall and seven-day record for group comparisons. J Am Diet Assoc 28:218–21. 1952.
10. Burke BS, Stuart HC. A method of diet analysis. J Pediatr 12:493–503. 1938.
11. Balogh M, Medalie JH, Smith H, Groen JJ. The development of a dietary questionnaire for an ischaemic heart disease survey. Isr J Med Sci 4:195–203. 1968.
12. McPherson RS, Nichaman MZ, Kohl HW, et. al. Intake and food sources of dietary fat among school children in The Woodlands, Texas. Pediatrics 86(4):520–6. 1990.
13. Block G, Hartman AM, Dresser CM, et al. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 124:453–69. 1986.
14. Lilienfeld AM, Lilienfeld DE. Foundations of epidemiology. New York: Oxford University Press Inc. 14,291. 1980.
15. Harlan WR, Hull AL, Schmouder RP, et al. Dietary intake and cardiovascular risk factors: Part II, serum urate, serum cholesterol, and correlates. National Center for Health Statistics. Vital Health Stat 11(227). 1983.
16. Nichols AB, Ravenscroft C, Lamphiear DE, Ostrander LD Jr. Independence of serum lipid levels and dietary habits: The Tecumseh Study. JAMA 236:1948–53. 1976.
17. Sobell J, Block G, Koslowe P, et al. Validation of a retrospective questionnaire assessing diet 10–15 years ago. Am J Epidemiol 130:173–87. 1989.
18. Jensen OM, Wahrendorf J, Rosenqvist A, et al. The reliability of questionnaire-derived historical dietary information and temporal stability of food habits in individuals. Am J Epidemiol 120:281–90. 1984.
19. Jain M, Howe GR, Johnson KC, et al. Evaluation of a diet history questionnaire for epidemiologic studies. Am J Epidemiol 111:212–9. 1980.
20. James WPT, Bingham SA, Cole TJ. Epidemiological assessment of dietary intake. Nutr Cancer 2:203–12. 1981.
21. Chalmers FW, Clayton MM, Gate LO, et al. The dietary record – How many and which days? J Am Diet Assoc 28:711–7. 1952.
22. Liu K, Stamler J, Dyer A, et al. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. J Chron Dis 31:399–418. 1978.
23. Sempos CT, Johnson NE, Smith EL, et al. Effects of intraindividual and interindividual variation in repeated dietary records. Am J Epidemiol 121:120–30. 1985.
24. Balogh M, Kahn H, Medalie JH. Random repeat 24-hour dietary recalls. Am J Clin Nutr 24:304–10. 1971.
25. Beaton GH. What do we think we are estimating? In: Beal VA, Laus MJ, eds. Proceedings of the Symposium on Dietary Data Collection, Analysis, and Significance, June 15–16, 1981. Massachusetts Agricultural Experiment Station, Research Bulletin No. 675. Amherst: University of Massachusetts. 36–49. 1982.
26. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65. 1985.
27. Jain MG, Harrison L, Howe GR, et al. Evaluation of a self-administered dietary questionnaire for use in a cohort study. Am J Clin Nutr 36:931–5. 1982.
28. Browe JH, Gofstein DM, Morlley DM, et al. Diet and heart disease study in the cardiovascular health center. J Am Diet Assoc 48:95–108. 1966.
29. Reshef A, Epstein LM. Reliability of a dietary questionnaire. Am J Clin Nutr 25:91–5. 1961.
30. Samet JM, Humble CB, Skipper BE. Alternatives in the collection and analysis of food frequency interview data. Am J Epidemiol 120:572–81. 1984.
31. Burke BS, Beal MA, Kirkwood SB, et al. Nutrition studies during pregnancy. Am J Obstet Gynecol 46:38–52. 1943.
32. Reed RB, Burke BS. Collection and analysis of dietary intake data. Am J Public Health 44:1015–26. 1954.
33. Jain M, Howe GR, Johnson KC, et al. Evaluation of a diet history questionnaire for epidemiologic studies. Am J Epidemiol 111:212–9. 1980.
34. Hankin JH, Reynolds WE, Margen S. A short dietary method for epidemiologic studies. II. Variability of measured nutrient intakes. Am J Clin Nutr 20:935–45. 1967.
35. Byers T, Marshall R, Fiedler R, et al. Assessing nutrient intake with an abbreviated dietary interview. Am J Epidemiol 122:41–50. 1985.
36. Van Horn L. Cardia Study. Personal communication. 1986.
37. Block G, Hartman AM, Naughton D. A reduced dietary questionnaire: Development and validation. Epidemiology 1:58–64. 1990.
38. Cummings R, Block G, McHenry K, et al. Evaluation of two food frequency methods of measuring dietary calcium intake. Am J Epidemiol 126:796–802. 1987.
39. Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. J Clin Epidemiol 43(12):1327–35. 1990.
40. Block G, Bowen P, Burgess M. Relationship between serum and self-reported dietary measures of carotenoids. Unpublished data.

Chapter 3 Dietary methods in cardiovascular disease critique

by Patricia J. Elmer, Ph.D., R.D., Division of Epidemiology, University of Minnesota

Introduction

Interest in the effects of nutrients and related dietary constituents on cardiovascular disease risk factors and outcomes is constantly increasing. An extensive body of literature exists, evaluating many nutrition and heart disease relationships. An enormous amount of knowledge is still to be gained about nutritional influences on the development and treatment of heart disease as well as the development of eating patterns. The relationship between dietary intake of fat and cholesterol and serum cholesterol and cardiovascular disease mortality has been extensively evaluated.

A myriad of questions related to dietary intake and risk factor development, levels of risk factors, and disease outcomes are currently under study. There is a great deal of interest in the interrelationships and interactions of various dietary intakes and other risk factors, health practices, and behaviors. Examples of current questions under study include: How does smoking affect nutrient intake, weight status, and disease outcome? How do smoking, alcohol, fat, caffeine, and exercise interact to affect serum lipids? What effect do specific fatty acids, monounsaturates, and omega three fatty acids have on serum lipids and coronary heart disease (CHD) mortality? What effect do dietary fat intake and total calorie intake have on body fat distribution? What effect do different minerals have on the etiology of hypertension? Are there particular ages at which individuals are particularly susceptible to dietary "insult" or dietary modification? How might the constellation of factors initiated during adolescence—smoking, alcohol use, use of birth control pills, other drug use, high fat diet, and the growth spurt—affect future risk factor development? Which individuals are sodium sensitive, and what effect do minerals other than sodium have on their blood pressure? What are the effects of weight loss, exercise, and fiber intake on serum lipids? What effect does dietary sodium intake have on sodium taste preference; do other minerals affect sodium taste preference? What effect do dietary interventions have on dietary intake and risk factor modification?

As can be seen from this very cursory list of current research questions, a tremendous diversity exists in the type and scope of investigations. Early work in nutrition and cardiovascular disease focused on a few macronutrients, particularly fat and cholesterol. The list of nutrients under investigation in the area of cardiovascular disease has lengthened considerably and includes not only calories, fat, and cholesterol but also carbohydrates, alcohol, caffeine, calcium, sodium, potassium, chloride, magnesium, and fiber. There is renewed interest in specific fatty acids rather than just the classification of individuals by saturated fat and polyunsaturated fat and there is interest in the effects of specific fiber components and specific food items on risk factors. It is likely that this list will continue to expand in the future, increasing the need for comprehensive data on nutrients and diet intake.

When choosing a method for the evaluation of dietary constituents and cardiovascular disease, a wide variety of factors need to be taken into consideration beyond which nutrients will be targeted for evaluation. What are the major questions to be answered through the data obtained? What is the nature of the dependent variables to be evaluated in relation to diet, and how are they to be measured—for example, mortality, risk factors, and physiological, behavioral, and psychological factors? What are the limitations of the dietary methodologies employed?

Many studies have failed to show correlations between nutrient intakes and various cardiovascular disease risk factors or outcomes within populations, although relationships have been seen between populations. A wide variety of factors have been suggested to contribute to this lack of findings (1). A large variability in the measures of the diet or risk factor may obscure the relationship; hence, crucial emphasis is put on using appropriate, valid, and reliable methodology. A very narrow range of nutrient intake in the population under study may exist. There may be a homogeneously high intake of the nutrient with a concomitant variation in the susceptibility to the nutrient, as is thought to be the case with sodium intake. Intraindividual variation in the nutrient intake or physiologic factor may be very high or even exceed the interindividual variation, resulting in misclassification of the individual intake or risk factor level. Are group or individual data needed to answer the question under study? There may be reporting bias in certain subgroups of the population, men versus women, young versus old, and so forth. There may be bias in the sampling or survey methods; days of the week may be excluded, potentially underestimating nutrients; or seasonal differences may exist in geographic areas and may not be reflected in the survey design. The disease effect (that is, the hyperlipidemic or hypertensiogenic effect) may begin in childhood or early adulthood, and current adult intake may not affect the factor under study; in this case, retrospective dietary information would be important. Changes in diet, other health practices, or medication may have occurred because of disease diagnosis or family history, thus obscuring the diet conditions that may have led to the disease. Major confounders could exist that have not been accounted for or measured. Examples include smoking, medication, exercise, and concurrent disease conditions. Such factors need to be addressed when choosing a dietary methodology and in the design and development of the overall survey so that specific research questions can be adequately evaluated.

Procedural and organizational questions also need to be addressed when choosing the dietary methodology. Is the method feasible given the time, staff, or funding available? Is appropriate staff available, and can unbiased interview techniques be developed? What is the nature of the coding and calculation process for determination of nutrients; are there adequate guidelines and coding rules? Do these procedures change over time? What differences are there between procedures from one survey to another? What is the nature of the food table (food codes and composition) to be used for calculation; what are the sources and vintage of the data; are the values complete for the nutrients under study?

A wide variety of methods have been used to estimate dietary intake in cardiovascular disease evaluations, including 24-hour recalls, food records, food histories, food frequencies, weighed diets, and observations. These methodologies have been reviewed and summarized extensively elsewhere. Given the time and logistics of the National Health and Nutrition Examination Survey (NHANES), weighed food intakes and observations of individuals' intakes would not be appropriate methodologies. The following will be a discussion of methods and research questions related to the assigned nutrients for cardiovascular disease, with recommendations for NHANES III.

Sodium and potassium

Excessive sodium and salt intake have been implicated in the development of hypertension in both animals and humans (2). Potassium intake is believed to have a potentially beneficial effect on blood pressure through a modifying effect on sodium. Dahl originally suggested the relationship between habitual salt intake in populations and the prevalence of hypertension (3). He showed that populations with low salt intakes have a prevalence of hypertension that is almost zero, compared with populations with high salt intakes, such as southern Japan, where the prevalence is very high. Glieberman further evaluated salt intake and blood pressure in other populations; he provided analyses of the regression of salt intake against diastolic blood pressure for 25 populations, showing a significantly greater average blood pressure in populations with greater salt intakes (4). In these studies, salt intake was estimated by several methods: 24-hour urine collections and estimates of dietary salt and sodium intake. These studies did not attempt to standardize blood pressure measurement, and the age groups included differed among populations. In addition, these studies did not account for other factors affecting blood pressure, such as weight, alcohol, exercise, levels of stress, or urbanization. Despite the limitations of such studies, all of which would tend to obscure any relationship, a relationship between blood pressure and sodium intake was established.

Sodium restriction in hypertensives has been shown to lower blood pressure (5–11). In a double-blind crossover trial of the effects of sodium ingestion in hypertensives, MacGregor et al. demonstrated a significant blood pressure lowering in the low-sodium condition after just 4 weeks of treatment (11). The effect of sodium intake and blood pressure in infants was studied using a double-blind design where newborn infants received formula and baby foods with either the normal amount of sodium or 50 percent less sodium for 6 months. After 25 weeks blood pressure was significantly different between the two groups, with a 2.1 millimeters of mercury difference in blood pressure between the regular and low-salt groups (10). Recent studies have also shown a blood-pressure-lowering effect with potassium supplementation in both hypertensive and normotensive groups. However, this effect is less well established and is being investigated further (9–13).

Lower blood pressures in Western vegetarian groups have also been related to high potassium intakes (14,15). Analyses of NHANES data show relationships between blood pressure and the nutrients sodium, potassium, alcohol, and calcium and between blood pressure and weight (16–18). Estimates of dietary sodium intake have been made by a variety of methods, including data from production and sale of table salt; estimates of salt and other sodium ingredients used in food manufacture; and data from dietary recalls and records and urine excretion (19–21). Many different sources of sodium contribute to the total daily intake of an individual. These include foodstuffs, water, and medications. Most foodstuffs have a very low naturally occurring sodium content. The majority of sodium contributed by foodstuffs comes from sodium that is added either in food manufacture (as sodium chloride or other sodium-containing flavoring or preservatives) or by the individual (as discretionary salt during cooking or at the table).

This addition of salt to the food supply in various ways and at various levels makes the task of quantifying dietary sodium intake extremely difficult. Although few prescription drugs contain sodium, many over-the-counter drugs contain significant amounts (22). For example, some antacids can contribute as much as 1,200 milligrams (mg) of sodium, and a single dose of aspirin can contain 50 mg of sodium. Estimates of sodium contributed by major categories are 10–20 percent from naturally occurring sodium, 40–60 percent from sodium added in food manufacture, 25–35 percent from "discretionary" salt, and the remainder from water and medications.

Market basket surveys conducted by the Food and Drug Administration during the period 1976–79 estimated available sodium content to be approximately 1,740 mg of sodium per 1,000 kilocalories (20). Food consumption and dietary survey data suggest that the total daily sodium intake ranges from 2,300 mg to 6,900 mg (10 to 14 grams of sodium chloride) per day. Males consume more sodium than females, and intake in both sexes declines with age. This decline is probably due to both a decline in calorie consumption and a reduction in salt use and consumption of high-sodium foods as a result of the diagnosis of hypertension or the concern about hypertension. Results of a recent statewide risk-factor telephone survey in Minnesota indicated that 30 percent of the adults responded that they did not use salt at the table. Age breakdown of the data showed that 65 percent of the adults ages 55 years and over did not use salt, compared with only 20 percent of those under age 55 years (23). Several dietary surveys indicate that children ages 9–12 years may consume the highest levels of sodium of all age groups. Data from NHANES I, based on 24-hour recall data, also showed that sodium intake varied with age and also by sex, with males ages 18–44 years and females ages 6–11 years consuming the maximum amounts (21). Food group analyses indicate that grain and cereal products are the largest contributor to sodium intake, followed by meat products, dairy foods, and fats and oils (20).

In the past 15 years a great deal of attention has been given to developing methods to assess sodium and potassium intake. The widely accepted standard for estimating sodium is the measurement of electrolytes by a 24-hour urine collection (24–29). It is believed that this method avoids the problems and biases related to collecting dietary data, such as subject memory, accuracy of portion sizes, coding, and data base concerns. However, even this standard method is an indirect method and subject to interpretation error (29). All sodium consumed in a 24-hour period is not necessarily excreted in that same period. Balance studies and comparison of analyzed duplicate portions of food compared with urine indicate that approximately 85–95 percent of the ingested sodium is excreted in the 24-hour urine collection. The 24-hour urine collections also suffer from the same problems of large intraindividual variability (in sodium intake and daily sodium excretion rate) that hamper dietary methods. Liu et al. estimate that as many as nine 24-hour urine collections are needed to characterize the individual intake (28,29). As is the case with dietary methods, investigators have attempted to develop "shortened" methods for estimating urine sodium and potassium excretion in epidemiologic studies. Overnight urine and spot urine collections as well as chloride titrator strips have been evaluated for this purpose (30,31). Generally good correlations were found between the 24-hour and overnight collections for sodium and potassium. Many believe that overnight urine collections can be used to characterize population means. When possible, multiple overnight collections are preferable to a single collection.

Nutrition surveys have often not included estimates of dietary sodium intake. The belief that it is impossible to determine sodium, in part because table and cooking salt cannot be estimated, and concerns over data base limitations in sodium values have limited the use of dietary sodium evaluations by some investigators. Early dietary studies in Japan, Korea, and Polynesia, each using a different dietary method—food history, weighed food intake calculated by food table values, and a 24-hour recall—showed good agreement between the dietary sodium values when compared with 24-hour urine sodium values (29). More recently a great deal of work has been conducted with dietary methods for the estimation of sodium and potassium. Schachter et al. compared 3-day food records and weighed food portions with 24-hour urine collections. Calculated food record sodium was 11 percent lower than the urine values, but the calculated potassium was only 1 percent lower than the urine potassium value (26). Caggiula et al. compared single- and multiple-day food records for their ability to estimate sodium, potassium, and energy intake and assessed the correspondence between sodium and potassium calculated from food records and urine analysis (32). A single-day food record approximated the 6-day food record mean closely for sodium, potassium, and calories. Sixty percent of the group could be classified correctly into quartiles based on a 1-day record compared with the 6-day mean. Correlation coefficients between the 1-day mean and the 6-day mean were only slightly lower than correlation coefficients between multiple days and the 6-day mean. Urine excretion of sodium and potassium were significantly correlated with the single- and multiple-day record values; correlations ranged from 0.50 to 0.76. Urinary sodium mean values were always higher, on average by 40 milliquivalents (meq), than the food record values, but the potassium differences between urine and food records were small. The value of the use of a single well-documented and calculated food record for estimation of group sodium and potassium intake was emphasized, but the authors pointed out that the estimation of individual intake would require multiple collections.

Elmer et al. evaluated the use of 24-hour recalls and 3-day food records compared with 24-hour urine collections for the estimation of sodium intake in middle-aged men who were participating in a sodium reduction trial (33). The food records were carefully collected and documented, salt use was estimated, and a special data base and calculation system were designed to capture sodium. At baseline there was excellent correspondence between the urine and 3-day food record average sodium values. The mean 24-hour urine sodium was 170 meq and the average 3-day food record intake was 169 meq. During the intervention period the 3-day food record underestimated the urine sodium levels by approximately 20 meq (20 percent). It is believed that part of this underestimate was due to participant desire to show good adherence to the low-sodium diet intervention advice. The 3-day food record and 24-hour recall were compared to see if one resulted in a better estimate of sodium intake. The sodium values for both the recalls and the 3-day food records were similar during all time periods. In order to compare the effects of the calculation and data base used for estimating dietary sodium, the 24-hour recalls were calculated using both the special sodium data base and the Nutrition Coding Center (NCC) data base, which did not accommodate sodium extensively at the time. When the recalls were calculated using the less sodium-sensitive NCC system, the recalls underestimated the 24-hour urine sodium levels by 33 percent. The conclusions are that carefully documented and calculated food records and 24-hour recalls could accurately predict 24-hour urine sodium values and that the calculation of salt added and the food table used were extremely important in improving the accuracy of the sodium estimates.

Simplified methods to estimate sodium intake have also been used. In a metabolic-ward setting, Frank et al. found good correlations between a self-administered short dietary inventory and urine sodium and potassium excretion (34). Pietinen, Tanskanen, and Tuomilehto (35) and Dahl (3) both classified individuals based on a few questions about their salt use habits. Pietinen et al. found this salt index to be significantly correlated with 24-hour urine sodium excretion. Dahl found classification by his salt index to be related to blood pressure, but this relationship was not replicated by others.

The recommended method for the estimation of sodium and potassium intake is by use of the 24-hour recall. Multiple nonconsecutive days should be collected to characterize the individual intake and to reduce the intraindividual variation. Both weekends and weekdays should be included. In addition, it is important to document and quantify salt and condiment use. Specifically designed interview probes for sodium and potassium should be employed. The food table utilized for calculation of nutrients needs to be comprehensive in its entries in order to adequately characterize sodium and potassium intake. Standardized coding rules for sodium capture should also be developed. For a case control study, information on previous use would also be recommended, as many individuals may have modified their intake based on the diagnosis or family history of hypertension. Urine collections for the characterization of sodium and potassium could be considered, utilizing either multiple 24-hour or overnight collections. The logistics of collecting these samples may be a potential barrier to obtaining these data, but substudies could be considered.

Calcium

There is substantive evidence linking cellular calcium metabolism and calcium intake with the pathophysiology of hypertension. Serum calcium has been positively correlated with blood pressure. Serum ionized calcium is reported to be lower in hypertensive persons than normotensive persons, and it has been reported to directly affect peripheral vascular tone (36,37). Calcium-regulating hormones may also affect blood pressure; parathyroid hormone can be stimulated by low dietary calcium, and elevated levels have been associated with hypertension in both humans and animals (37). Several recent investigations have demonstrated associations between dietary calcium intake and blood pressure. The observation that regions with hard water had lower cardiovascular mortality initiated the original epidemiologic interest in relationships between calcium and cardiovascular disease. An inverse relationship between dietary calcium intake and blood pressure has been reported for several population groups (37). Cross-cultural comparisons of the prevalence of gestational hypertension have shown that countries with higher calcium intakes have lower prevalence. However, these studies did not take into account many other variables that may have been related to the development, detection, and reporting of gestational hypertension. Recent evidence also suggests a hypotensive effect of calcium supplementation in some hypertensive persons (38,39).

Several dietary surveys suggest that hypertensive persons as a group consume less calcium than normotensive persons. Langford et al. report that, based on their surveys, rural black hypertensive women consume diets very low in calcium (24). In Oregon, 24-hour dietary recalls were used to assess dietary intake of a small group of hypertensive and normotensive persons. Hypertensive persons consumed, on the average, 25 percent less calcium than normotensive persons (40). Cheese consumption was the major food item of difference, with normotensive persons consuming larger amounts. It is not clear if this difference in food consumption was a longstanding pattern or a recent phenomenon caused by concern over hypertension. The difference in cheese consumption may have been due to an attempt to reduce sodium or fat intake by a hypertensive group trying to change their diet.

In a survey of 2,326 men in southern California, a single question was asked on the usual number of cups of whole milk consumed daily. Whole milk consumption was significantly lower for borderline, untreated, and treated hypertensive men; but among women, consumption was lower only for treated hypertensive women (41). A 22-percent sample of this population also had a 24-hour diet recall. Only dairy calcium intake was calculated. Lower dairy calcium consumption was found in untreated hypertensive, but not borderline hypertensive, persons. Total dairy calcium intake adjusted for age, obesity, and alcohol intake was correlated with diastolic blood pressure. Lower calcium intakes in untreated hypertensive persons were also reported in analyses of the NHANES I data which utilized 24-hour diet recalls to quantify calcium intake (16–18). Calcium intake was higher for men than women, but the ratio of calcium per 1,000 calories was the same for men and women. Calcium intake was lower in the older age groups for both men and women.

As is the case with other nutrients, a single measure of current calcium intake may not provide the most relevant information related to cardiovascular disease risk factors or outcomes. This is also true for the relationship of dietary calcium to bone mineralization and osteoporosis, which will be reviewed elsewhere in this workshop. Multiple, nonconsecutive, 24-hour recalls can be utilized to provide an estimate of intake over a 1-year period of time. Use of 24-hour recalls will maintain comparability with previous NHANES and will provide precise information so that individuals can be ranked based on their calcium intake. Multiple days of collection will reduce the intraindividual variability. If long-term calcium intake is needed for evaluation of a specific question, a food history and retrospective information would be required. It is important to note that even though many hypertension studies are focusing on calcium intake, a great many nutrients are likely to be involved in the development of elevated blood pressure, including sodium, potassium, chloride, magnesium, vitamin D, fatty acids, and alcohol. The effect of calcium administration on blood pressure lowering may be different in different population groups and may vary by other dietary conditions, such as sodium and potassium levels. Collection of dietary data that merely focus on sources of calcium would be inadequate for the evaluation of nutrient and blood pressure relationships.

Alcohol

The associations of alcoholic beverage consumption with heart and other cardiovascular diseases are complex. Alcohol has diverse effects on the endocrine, gastrointestinal, metabolic, hematopoietic, and neurologic systems that can affect the heart and circulatory system. Several epidemiologic investigations have demonstrated that alcohol intake is positively associated with systolic and diastolic blood pressure and the incidence of hypertension (16,17,42–45). However, data also indicate that regular moderate use of alcohol is associated with a lower risk of major coronary events (42,43). High-density lipoproteins (HDL's) have been suggested as possible mediating factors for this reported protective effect of alcohol consumption (46). Heavy alcohol consumption and problem drinking are associated with an increased mortality from all causes, cardiovascular disease, CHD, and cancer (42–47).

Problems have arisen in comparing studies of the relationship of alcohol with disease because of differences in the type of alcohol measure used to evaluate consumption. Measures fall into two distinct groups: Socially defined measures (alcoholic or problem drinkers based on behavior patterns) and quantity-frequency measures. A few studies have used medical diagnosis of alcoholism or liver cirrhosis to classify individual alcohol intake (48). Smoking status, other dietary components, socioeconomic status, psychological stress, and personality type have been suggested as potential major confounders in the determination of associations between alcohol and both cardiovascular risk factors and disease outcomes. Potential misclassification of individuals into consumption groups (particularly "ever drink" versus "never drink") may also contribute to the findings that abstinence is associated with higher mortality than light or moderate intake.

Alcohol consumption varies by age, sex, socioeconomic status, ethnicity, and religious affiliation. More men than women drink alcohol, and more men report heavy consumption. A greater percentage of older adults than younger adults report no drinking. Approximately 30 percent of all adults classify themselves as nondrinkers and, using various survey methods, approximately 9 percent of adults are classified as chronic heavy drinkers. Respondents who are obese, who smoke, and who do not use seatbelts are more likely to be chronic heavy alcohol consumers (16,49).

The Lipid Research Clinics Prevalence Study assessed alcohol intake in children and adolescents by questionnaire using the past week as the consumption reference time period. Twenty-two percent of the respondents ages 12–18 years reported consumption in the past week, with the prevalence increasing from 6 percent for ages 12–13 years to 53 percent for ages 18–19 years (50). The majority of investigations that have evaluated the relationship between alcohol consumption and cardiovascular disease have utilized some type of alcohol intake history for obtaining data on the quantity and frequency of intake over some specified period of time. The time period of reference has varied from 1 week to 1 year. The average number of alcoholic drinks per day and the absolute amount of ethanol are the two most commonly reported variables of consumption (41–50). Analyses often classify individuals into consumption categories or patterns: Abstainer, light, moderate, heavy, chronic heavy, or alcoholic. These classifications can be extremely important in the elucidation of the alcohol-disease relationships. It is also important to consider not only current classification but past classification. Several investigations have shown that light to moderate consumers of alcohol have lower mortality rates than either heavy drinkers or abstainers; but when the data are analyzed separately for abstainers who never drank versus abstainers who were former drinkers, the former drinkers accounted for most of the deaths in the aggregate abstainer groups (43).

Currently there is interest in characterizing drinking patterns further to encompass the concept of "binge drinking," cases in which the usual frequency or the usual average amount consumed may be low to moderate, but a pattern of excessive consumption exists with some frequency (for example, no alcohol during the week but 7–10 drinks on one or more weekend days). This pattern could yield an average weekly consumption of 1–4 drinks per day.

Validity and reliability are major research issues related to self-reported alcohol intake data. A variety of methods have been used to evaluate the reliability and validity of self-reported alcohol intake. It is generally believed that alcohol intake is underreported and that the underreporting is not random (51). Several investigations indicate that heavy alcohol consumers or alcoholics markedly underreport intake. Denial of an alcohol problem, cultural norms, concern over being classified as alcoholic, and inability to recall alcohol intake during heavy drinking bouts have been cited as reasons for underreporting (51,52). Other investigations suggest that part of the underreporting might be related to the selection options that respondents are given as categories of consumption on questionnaires of alcohol intake. These questionnaires usually focus on low frequencies and low quantities. Data suggest that the majority of alcoholic beverages are consumed by a small group of heavy consumers. It is estimated that the heaviest drinking 10 percent of the population may consume as much as 40–50 percent of all alcoholic beverages. Providing questions that focused on heavy consumption yielded higher reported intakes when two questionnaires were compared in an investigation of alcoholics' self-reported intake (53). The reliability of general population surveys of consumption has been investigated by using alternate forms, test-retest methods, and 30-day diaries for comparisons. Findings indicated a higher level of reliability of reported alcohol intake for the general population than for the consumption measures used for the alcoholic population (54). Klatsky and Friedman also point out that underreporting would have the effect of dilution of the ability to see an alcohol-disease association, but given the underreporting, associations are still seen (43).

The recommended method for collection of alcohol intake data is by use of a quantitative beverage frequency-history technique, which would cover a minimum of 1 week's usual intake and would include questions on maximum amounts used and previous consumption status. Difficulty in quantifying amounts of alcohol consumed may in part be due to the large number and types of beverage containers and the fact that many individuals may not actually pour or prepare their own alcoholic beverages and hence not see the amount of liquor used. The use of portion-size aids could help to facilitate estimation of consumption. An additional problem related to alcohol intake evaluation is the potential effect of acute withdrawal from alcohol. Some investigators suggested that blood pressure may rise during acute withdrawal; the nondrinking period that the respondent is spending in the survey center or clinic in some cases could be considered in this light. The respondent may report no alcohol consumption in the past 24 hours, and hence be classified as a low, moderate, or nondrinker, and yet be experiencing the blood pressure effects of acute withdrawal. Attempts to classify usual patterns of alcohol consumption and binge drinking may alleviate some of these difficulties.

Caffeine

Caffeine is found naturally in coffee, tea, and cocoa; as an additive in other beverages; and in such medications as antacids, analgesics, and weight loss aids. Coffee is the greatest dietary source of caffeine in the United States, providing approximately 75 percent of the population intake. Tea provides 15 percent, and soft drinks, cocoa, and medications, the remainder (55). Age, sex, and geographic differences exist in the patterns of caffeine consumption. These patterns change with age; the majority of caffeine for children and adolescents comes from soft drinks, and coffee is the major source for adults to about age 60, when many individuals switch to tea. Asian, Hispanic American, and black persons report consuming lower amounts of coffee than white persons report, although it is not clear if the total caffeine consumption is lower (56). In a report on a survey of substances Generally Recognized as Safe (GRAS), the distribution of daily caffeine consumption was estimated. Fourteen percent of infants 2 years of age consumed some caffeine during a 2-week period. For individuals ages 18 years and over, 82 percent consumed caffeine, with the mean intake being 186 mg per day, a little over the equivalent of two cups of coffee. Seventy-four percent of the pregnant women surveyed consumed caffeine, with a similar mean intake of two cups per day (57). Caffeine and coffee consumption are not routinely reported as part of many general nutrition and food consumption surveys.

An evaluation of the literature on relationships between caffeine consumption and cardiovascular disease reveals that most investigators report coffee, not caffeine, consumption. In some instances, only data on coffee intake were obtained; in other cases, no relationship was found with beverages other than coffee.

Several studies have shown a link between increased coffee consumption and an increased incidence of heart disease (58–61); other investigations have failed to show this relationship (62–65). The pathophysiologic mechanisms linking caffeine to cardiovascular disease are not clear, although several hypotheses have been suggested. The possibility that coffee may produce ventricular arrhythmias has been suggested (66). Associations between coffee consumption and serum lipids and lipoproteins have been demonstrated by several investigators (67–69). However, many of the positive associations between caffeine and cardiovascular disease have been attributed to the confounding effects of age, smoking, obesity, and alcohol consumption. Other dietary factors, such as a high-fat diet, may also confound the coffee and serum lipoprotein association (66).

The majority of the reports relating caffeine to heart disease have utilized beverage frequency or a history technique to obtain consumption data (58–60). Cups of coffee consumed is the most commonly employed variable for analysis; milligrams of caffeine may also be calculated. To date no validation of different methods of obtaining caffeine intake, such as comparisons between diet histories and frequencies or weighed or measured portions, have been reported; hence, there are little published data to evaluate the validity or reliability of self-reported methods.

When a nutrient or a dietary component has a limited distribution in the food supply, as caffeine does, or the consumption patterns are very different for different populations, focused questions about the frequency and quantitation of intake are more useful, in general, than a single-day intake from a record or recall. The use of a quantitative food frequency record would be recommended in this case; it would be appropriate for the case control, cohort, and cross-sectional designs. The major problem associated with collection and quantification of caffeine intake data is distinguishing between caffeinated and decaffeinated items. The ability of the respondent to accurately quantify portions and the choice of the caffeine values to be used in calculation are other potential problems. Use of portion-size aids and probes on caffeination should alleviate some of these respondent difficulties. Because caffeine has potent pharmokinetic properties, long-term history of use has been of interest in some investigations of all causes of mortality in addition to cardiovascular disease and cancer. Some previous studies have been criticized for the failure to collect information on the duration of exposure to caffeine (56).

Fat, calories, and cholesterol

Observations of dramatic differences in CHD mortality between geographic areas raised questions about the possible associations of mortality with differing dietary habits. Since the 1940's, the role of nutrition in the etiology of CHD has been extensively studied in a variety of epidemiologic settings. Calories, fat, and cholesterol were the focus in many of the early investigations, and they continue to be of importance in current studies (70). These nutrients have been determined by a variety of methods (24-hour recalls, food records, diet histories, food frequencies, weighed intakes, and aggregate national food production and disappearance data). Various prospective epidemiologic studies have demonstrated significant correlations between dietary fats, especially saturated fats, and serum cholesterol and CHD mortality. The Seven Countries Study, using weighed food intake for multiple days, demonstrated correlations of $r = 0.87$ for percent of calories and saturated fatty acids with serum cholesterol and $r = 0.84$ with CHD mortality (71). The Ni-Hon-San study of men of Japanese descent in Hawaii, California, and Japan utilized the 24-hour recall (72). Differences in CHD incidence paralleled the marked difference in fat consumption. Residents in Japan had the lowest risk and those in California, the highest. The Western Electric Study (73) and the Ireland-Boston Diet Heart Study (74) utilized the diet history method. In both studies, dietary cholesterol consumption, saturated fatty acid intake, and the Keys dietary score significantly predicted subsequent 20-year CHD mortality. In a 20-year followup of Seventh Day Adventists, using focused diet frequency questions, it was observed that for both men and women, meat eating was associated with an increased risk of CHD and risk increased as frequency of meat consumption increased (75). In the Zuthphen study, 20-year CHD was also observed prospectively (76). Dietary intake was estimated using a food history. The investigators found an inverse association of fish consumption with CHD mortality. Dietary cholesterol intake was also a significant predictor of CHD death.

It is generally well accepted that population comparisons demonstrate strong associations for dietary cholesterol and saturated fat with CHD mortality and serum lipids. Within-population comparisons have generally showed smaller or no correlations. The homogeneity of the diet and inadequacies of the dietary methodology to differentiate correctly an individual's fat intake have been cited as reasons for the lack of association (77–79). Several studies that found no association utilized 24-hour recalls. Day-to-day variability in dietary fat intake within an individual is larger than between individuals (79). A single day's intake is not adequate to characterize an individual's fat, cholesterol, or calorie intake because of this large intraindividual variability (79,80). However, it should be noted that some prospective studies, utilizing diet histories, have also failed to demonstrate associations between dietary fats and CHD mortality. Use of a diet history method does not guarantee demonstration of association between a nutrient and disease or risk factor. Design and methodologic protocol are extremely important. In many reports the methodologies are not well described. There is very little mention of how food items are coded and what type of data bases were utilized for calculation, what the interview procedures were, how long the interview lasted, and so forth. More recently, standardization of interviews, coding procedures, special methods to estimate fat, and more extensive food tables have enhanced the ability to estimate fat intake across all methods.

Extensive work has been conducted to evaluate the validity and reliability of the various dietary intake measures (81–91). Generally the diet history tends to produce higher estimates for all nutrients, including fat, cholesterol, and calories. When diet histories have been repeated, good agreement has been found between the repeat measures (77). The diet history is considered by many to provide better information on the individual's usual intake and does not suffer from the problems of large intraindividual variation (81–83). However, if the history is collected for a long period of time, such as 1 year, the respondent may have considerable difficulty recalling intake. If the time period of collection is 1 month, seasonal variation is still not accounted for. Weighed food portions have been compared with 24-hour recalls, with good agreement except for calories (90). Studies of 24-hour diet recalls demonstrate the considerable variability of individual intakes of fat, calories, and cholesterol. It has been estimated that from 4 to 15 recalls would be needed to classify an individual within 5 to 20 percent of his or her true mean for calories, fat, and cholesterol (77,80,88,91).

There has been considerable interest recently in the development of quantitative food frequencies and item-based diet histories for epidemiologic and clinical studies (84,86,92,93). In a recent report on the validity of a semiquantitative food frequency, Willett et al. found good correlations between food record values and the questionnaire values for cholesterol, fat, saturated fat, polyunsaturated fat, and total calories (86). The food records provided slightly higher values for these nutrients than did the food frequencies. He found that the frequency questionnaire based on an abbreviated list of foods was able to classify individuals into quartiles of intake with good correspondence to the food records. Byers et al. used an abbreviated interview based on 128 foods and compared intake values to a more extensive food history with good results for fats and calories; classification by quintiles showed good agreement with the longer method (84). These studies indicate that, at least in some populations, use of shortened methods to estimate fat, cholesterol, and calories can provide valid and reliable results. The method of coding and calculating the nutrients, fat, calories, and cholesterol can produce differing results based on the procedures and data base used. It is important to utilize standardized methods and have specific rules as well as a current data base that contains manufactured food product data so as to adequately reflect the fatty acid composition of the diet and capture fat and calories that could be omitted without standardized procedures (94).

The continued theme of the review of dietary methods is that a single day's intake is inadequate to characterize correctly an individual's usual intake and to allow the correct placement of that individual into some range of the nutrient distribution (such as quintiles). Multiple days of diet information will be needed for nutrients, fat, cholesterol, and calories if individual comparisons are to be made for nutrient and disease or risk factor associations. Multiple 24-hour recalls would provide accurate data and maintain comparability with the previous NHANES. The days should be randomly selected and nonconsecutive. This type of methodology would necessitate telephone administration. Information on usual types of fats used, such as soft margarines and olive oil, and consumption of fish would be useful for characterizing specific fatty acid intake patterns. If multiple 24-hour recalls cannot be completed, additional food history questions should be utilized.

Summary

It is planned that a wide variety of questions related to nutrient intake and cardiovascular disease will be addressed in NHANES III. Given the scope of these questions and the number of different types of designs that may be utilized—cross-sectional, cohort, and case control—several different dietary methodologies will probably be needed within NHANES III. Assessment and ranking of current intake is a primary objective to the overall NHANES III design. The 24-hour recall will provide this assessment and maintain comparability with the previous extensive NHANES.

In many instances, however, information on usual intake or past intake will be necessary to better elucidate a nutrient-disease relationship, either because of temporal exposure concerns or because of the variability of the nutrient intake. A food history or quantitative frequency method for the specific nutrient(s) would best address these needs. Information on specific foods or other diet constituents (such as fish, dairy foods, alcohol, or water intake) in addition to nutrient levels will be of interest not only for cardiovascular disease but, with some food items, for cancer and osteoporosis as well. A history or quantitative frequency method would be appropriate for this purpose.

The question of serial and multiple measures needs to be addressed. For adequate classification of individuals, several nonconsecutive measures would be most appropriate. The actual number of measures would be determined by the estimate of the most acceptable reduction in variance that could be achieved with multiple measures, balanced by the feasibility and cost of such measures. The use of multiple measures would necessitate the use of alternative interview techniques, most likely by telephone. It is clear that, if this type of methodology is adopted, further validation and extensive pretesting would need to be conducted. A subsample approach could also be utilized.

Automated coding of dietary data and computer-prompted immediate data entry interview techniques may facilitate the use of these dietary data collection methods. Whichever methods are utilized, standardization of the coding procedures and additions to the nutrient data base, particularly for sodium, are essential.

Finally, there is considerable interest in evaluating any deleterious effects of dietary modification on health and mortality outcome. For example, there is concern that a fat-modified diet may increase cancer mortality in some or that a low-sodium diet may be unpalatable and lead to a lowered dietary intake and potential nutrient deficiencies. It would be useful in this regard for NHANES III to be able to ascertain if altered dietary levels were due to self-selected dietary practices, to prescription of a therapeutic diet, or to a disease state or illness that lowered overall food intake.

References

1. Prineas R, Blackburn H. Clinical and epidemiological relationships between electrolytes and hypertension. In: Horan M, Blaustein M, Dunbar J, et al., eds. NIH Workshop on Nutrition and Hypertension. New York: Biomedical Information Corp. 1985.
2. Porter G. Chronology of the sodium hypothesis in hypertension. Ann Intern Med 98(5, pt 2):720–4. 1983.
3. Dahl L. Salt intake in hypertension. In: Genset J, ed. Hypertension. New York: McGraw-Hill Book Co., 548–58. 1977.
4. Glieberman L. Blood pressure and dietary salt in human populations. Ecol Food Nutr 2:143–6. 1973.
5. Parijs J, Joossens J, Linden L, et al. Moderate sodium restriction and diuretics in the treatment of hypertension. Am Heart J 85:22–34. 1973.
6. Carney S, Morgan T, Wilson M, et al. Sodium restriction and thiazide diuretics in the treatment of hypertension. Med J Aust 1:803–7. 1975.
7. Morgan T, Adam W, Gilles A, et al. Hypertension treated by salt restriction. Lancet 1:227–30. 1978.

8. Fagerberg B, Anderson O, Isaksson B, et al. Blood pressure control during weight reduction in obese hypertensive men: Separate effects of sodium and energy restriction. Br Med J 288:11–14. 1984.
9. Parfrey P, Goodwin F, Evans S. Blood pressure and hormonal changes following alteration in dietary sodium and potassium in mild hypertension. Lancet 1:56–63. 1981.
10. Hofman A, Hazebroek A, Valkenburg H. A randomized trial of sodium intake and blood pressure in newborn infants. JAMA 250:370–3. 1983.
11. MacGregor G, Markandu N, Smith S, et al. Moderate potassium supplementation in essential hypertension. Lancet 2:567–70. 1982.
12. Khaw K, Thom S. Randomized double blind cross-over trial of potassium on blood pressure in normal subjects. Lancet 2:1127–9. 1982.
13. Khaw K, Barrett-Conner E. Dietary potassium and blood pressure in a population. Am J Clin Nutr 39:963–8. 1984.
14. Ophir O, Peer G, Gilad J, et al. Low blood pressure in vegetarians: The possible role of potassium. Am J Clin Nutr 37:755–62. 1983.
15. Meneely G, Battarbee H. High sodium-low potassium environment and hypertension. Am J Cardiol 3:318–26. 1976.
16. Gruchow H, Sobocinski K, Barboriak J. Alcohol, nutrient intake, and hypertension in U.S. adults. JAMA 253(11):1567–70. 1985.
17. Harlan W, Hull A, Schmouder R, et al. Blood pressure and nutrition in adults. Am J Epidemiol 120:17–28. 1985.
18. McCarron D, Morris C, Henry H. Blood pressure and nutrient intake in the United States. Science 224:1392–8. 1984.
19. Dresser CM, Carroll MD, Abraham S. Selected findings: Food consumption profiles of white and black persons 1–74 years of age in the United States, 1971–74. Advance data from vital and health statistics; no 21. Hyattsville, Maryland: National Center for Health Statistics. 1978.
20. Shank FR. Recent data on the amounts of sodium and sodium being consumed. In: White PL, Crocco SC, eds. Sodium and potassium in food and drugs. Chicago: American Medical Association. 1980.
21. Abraham S, Carroll MD. Fats, cholesterol, and sodium intake in the diet of persons 1–74 years: United States. Advance data from vital and health statistics; no 54. Hyattsville, Maryland: National Center for Health Statistics. 1979.
22. Fregly M. Attempts to estimate sodium intake in humans. In: Horan M, Blaustein M, Dunbar J, et al., eds. NIH Workshop on Nutrition and Hypertension. New York: Biomedical Information Corporation. 1985.
23. Healthy people. The Minnesota Experience MDH Center for Health Statistics. 1983.
24. Watson R, Langford H, Abernethy J, et al. Urinary electrolytes, body weight, and blood pressure: Pooled cross sectional results among four groups of adolescent females. Hypertension 2(suppl I):93–8. 1980.
25. Pietinen PI, Findley TW, Clausen JD, et al. Studies in community nutrition: Estimation of sodium output. Prev Med 5:400–7. 1976.
26. Schachter J, Harper PH, Radin ME, et al. Comparison of sodium and potassium intake with excretion. Hypertension 2:695–9. 1980.
27. Holbrook JT, Patterson KY, Bodner JE, et al. Sodium and potassium intake and balance in adults consuming self-selected diets. Am J Clin Nutr 40:786–93. 1984.
28. Liu K, Cooper R, McKeever J, et al. Assessment of the association between habitual salt intake and high blood pressure: Methodologic problems. Am J Epidemiol 110:219–26. 1979.
29. Liu K, Cooper R, Soltero I, Stamler J. Variability in 24-hour urine sodium excretion in children. Hypertension 1:631–6. 1979.
30. Liu K, Dyer AR, Cooper R, et al. Can overnight urine replace 24-hour urine collection to assess salt intake? Hypertension 1:529–35. 1979.
31. Watson R, Langford H. Usefulness of overnight urines in population groups. Am J Clin Nutr 23(3):290–304. 1970.
32. Caggiula AW, Wing RR, Nowalk MP, et al. The measurement of sodium and potassium intake. Am J Clin Nutr 42:391–8. 1985.
33. Elmer PJ, Jacobs DR, Gillum R, Prineas RP. Methods for accurate assessment of dietary sodium intake in mild hypertension. Circulation 9:107. 1983.
34. Frank G, Nicolich J, Voors AW, et al. A simplified inventory method for quantifying dietary sodium, potassium and energy. Am J Clin Nutr 38:474–80. 1983.
35. Pietinen P, Tanskanen A, Tuomilehto J. Assessment of sodium intake by a short dietary questionnaire. Scand J Soc Med 10:105–12. 1982.
36. McCarron D, Morris C. Calcium and hypertension. In: Horan M, Blaustein M, Dunbar J, et al. eds. NIH Workshop on Nutrition and Hypertension. New York: Biomedical Information Corporation. 1985.
37. Henry H, Morris C, Parrott-Garcia M. Increasing calcium intake lowers blood pressure: The literature reviewed. J Am Diet Assoc 85:182–5. 1985.
38. Johnson NE, Smith E, Freudenheim J. Effects on blood pressure of calcium supplementation of women. Am J Clin Nutr 42:12–7. 1985.
39. McCarron D, Morris C. Blood pressure response to oral calcium in persons with mild to moderate hypertension. Ann Intern Med 103:825–31. 1985.
40. McCarron D, Morris C, Cole C. Dietary calcium in human hypertension. Science 217:267–9. 1982.
41. Barrett-Conner E, Suarez L. Dairy products, calcium and blood pressure. Am J Clin Nutr 38:457–61. 1983.
42. Kagan A, Yano K, Rhoads G, McGee D. Alcohol and cardiovascular disease: The Hawaiian experience. Circulation 64(2):27–31, monograph 81. 1981.
43. Klatsky A, Friedman G. Alcohol and hypertension. In: Horan M, Blaustein M, Dunbar J, et al., eds. NIH Workshop on Nutrition and Hypertension. New York: Biomedical Information Corp. 1985.
44. Dyer A, Stamler J, Paul O, et al. Alcohol, cardiovascular risk factors and mortality: The Chicago experience. Circulation 64(2):20–30, monograph 81. 1981.
45. Wallace R, Lynch C, Pomrehn M, et al. Alcohol and hypertension: Epidemiologic and experimental considerations. The Lipid Research Clinics Program. Circulation 64(2):41–7, monograph 81. 1981.
46. Gordon T, Ernst N, Fisher M, Rifkind B. Alcohol and high density lipoprotein cholesterol. Circulation 64(2):63–6, monograph 81. 1981.
47. Cahalan D. Quantifying alcohol consumption: Patterns and problems. Circulation 64(2):7–13, monograph 81. 1981.

48. Celentano D, Martinez RM, McQueen D. The association of alcohol and hypertension. Prev Med 10:590–602. 1981.
49. Gentry E, Kalsbeek W, Hogelin G, et al. The risk factor surveys. II. Design, methods, and estimates from combined state data. Am J Prev Med 1(6):9–14. 1985.
50. Glueck C, Heiss G, Morrison J, et al. Alcohol intake, cigarette smoking and plasma lipids and lipoproteins in 12–19 year old children: The Collaborative Lipid Research Clinics Study. Circulation 64(2):48–56, monograph 81. 1981.
51. Watson C, Tilleskjor C, Hoodeshedk-Schow E, et al. Do alcoholics give valid self-reports? J Stud Alcohol 45(4):344–8. 1984.
52. Pernanen K. Validity of survey data on alcohol use. In: Gibbins R, Israel Y, Kalant H, eds. Research advances in alcohol and drug problems, vol. I. New York: John Wiley & Sons, Inc. 1974.
53. Poikolainen K, Karkkainen P. Nature of questionnaire options affects estimates of alcohol intake. J Stud Alcohol 46(3):219–22. 1985.
54. Williams G, Aitken S, Malin H. Reliability of self-reported alcohol consumption in a general population survey. J Stud Alcohol 46(3):223–7. 1985.
55. Roberts H, Barone J. Caffeine: History and use. Food Tech 9:32–9. 1983.
56. Pozniak P. The carcinogenicity of caffeine and coffee: A review. J Am Diet Assoc 85:1127–33. 1985.
57. Graham D. Caffeine—its identity, dietary sources, intake and biological effects. Nutr Rev 36:97–102. 1978.
58. Jick H, Miettinen OS, Neff RK, et al. Coffee and myocardial infarction. N Engl J Med 289:63–7. 1973.
59. Wilhelmsen L, Tibblin G, Elmfeldt D, et al. Coffee consumption and coronary heart disease in middle-aged Swedish men. Acta Med Scand 201:547–52. 1977.
60. Report from the Boston Collaborative Drug Surveillance Program: Coffee drinking and acute myocardial infarction. Lancet 2:1278–81. 1972.
61. LaCroix A, Mead L, Liang K, Pearson T. Coffee consumption and 25 year incidence of coronary heart disease: Effects of coffee drinking and smoking habits during adulthood. Circulation 72(pt 2):53. 1985.
62. Klatsky A, Friedman G, Sieglaub A. Coffee drinking prior to acute myocardial infarction: Results from the Kaiser-Permanente epidemiologic study of myocardial infarction. JAMA 226:540–3. 1973.
63. Dawber T, Kannel W, Gordon T. Coffee and cardiovascular disease: Observations from the Framingham study. N Engl J Med 291:871–4. 1974.
64. Yano K, Rhoads G, Kagan A. Coffee, alcohol and risk of coronary heart disease among Japanese men living in Hawaii. N Engl J Med 297:405–9. 1977.
65. Rosenberg L, Slone D, Shapiro S, et al. Coffee drinking and myocardial infarction in young women. Am J Epidemiol 111:675–81. 1980.
66. Prineas R, Jacobs D, Crow R, et al. Coffee, tea and VPB. J Chron Dis 33:67–72. 1980.
67. Thelle DS, Aneson E, Forde DH. The Tromso Heart Study: Does coffee raise serum cholesterol? N Engl J Med 308:1454–7. 1983.
68. Hafner S, Knapp A, Stern M, et al. Coffee consumption, diet and lipids. Am J Epidemiol 122:1–11. 1985.
69. Williams P, Wood P, Vranizan K, et al. Coffee intake and elevated cholesterol and apolipoprotein B levels in men. JAMA 253:1407–11. 1985.
70. Stamler J. Population studies. In: Levy R, Rifkind B, Dennis B, Ernst N, eds. Nutrition lipids and coronary heart disease. New York: Raven Press, 25–88. 1979.
71. Keys A. Seven countries: A multivariate analysis of death and coronary heart disease. Cambridge, Massachusetts: Harvard University Press. 1980.
72. Kato H, Tillotson J, Nichaman M, et al. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California. Am J Epidemiol 27:345–64. 1973.
73. Shekelle RB, Shryock AM, Paul O, et al. Diet, serum lipids and death from coronary heart disease: The Western Electric Study. N Engl J Med 304:65–70. 1981.
74. Kushi LH, Lew RA, Stare FJ, et al. Diet and 20-year mortality from coronary heart disease: The Ireland-Boston Diet Heart Study. N Engl J Med 312:811–8. 1985.
75. Snowdon DA, Phillips RL, Fraser GE. Meat consumption and fatal ischemic heart disease. Prev Med 13:490–500. 1984.
76. Kromhout D, Bosschieter EB, Coulander CD. The inverse relationship between fish consumption and 20 year mortality from coronary heart disease. N Engl J Med 312:1205–9. 1985.
77. Liu K, Stamler J, Dyer A, et al. Statistical methods to assess and minimize the role of intra-individual variability in observing the relationship between dietary and serum cholesterol. J Chron Dis 31:399–418. 1978.
78. Jacobs DR, Anderson JT, Blackburn H. Diet and serum cholesterol: Do zero correlations negate the relationship? Am J Epidemiol 110:77–87. 1979.
79. Beaton GH, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J Clin Nutr 32:2546–59. 1979.
80. Sempos CT, Johnson NE, Smith EL, Gilligan C. Effects of intra-individual and inter-individual variation in repeated diet records. Am J Epidemiol 121:120–30. 1985.
81. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 115(4):492–505. 1982.
82. Fehily A. Epidemiology for nutrition: Survey methodology. Hum Nutr Appl Nutr 37A:419–25. 1983.
83. Lyon J, Gardner J, West D, Mahoney A. Methodological issues in epidemiologic studies of diet and cancer. Cancer Res 43(suppl):2392s–96s. 1983.
84. Byers T, Marshall J, Fiedler R, et al. Assessing nutrient intake with an abbreviated dietary interview. Am J Epidemiol 122(1):41–50. 1985.
85. Mullen BJ, Krantzler NJ, Grivetti LE, et al. Validity of a food frequency questionnaire for the determination of individual food intake. Am J Clin Nutr 39:136–43. 1984.
86. Willett W, Sampson L, Stampfer M, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65. 1985.
87. Hankin J, Messinger H, Stallones R. A short dietary method for epidemiological studies. IV: Evaluation of a questionnaire. Am J Epidemiol 91:562–7. 1970.
88. Graham S. Diet and cancer. Am J Epidemiol 112:247–52. 1980.
89. Marshal J, Priore R, Haughey B, et al. Spouse-subject interviews and the reliability of diet studies. Am J Epidemiol 112:675–83. 1980.
90. Sorenson A, Calkins B, Connolly M, et al. Comparison of nutrient intake determined by four dietary intake methods. J Nutr Ed 17:92–9. 1985.

91. Balogh M, Medalie JH, Smith H, Groen JJ. The development of a dietary questionnaire for an ischemic heart disease survey. Isr J Med Sci 4:195–203. 1968.
92. Block G, Dresser C, Hartman A, Carroll M. Nutrient sources in the American diet: Quantitative data from the NHANES II survey. II. Macronutrients and fat. Am J Epidemiol 122:27–40. 1985.
93. Block G, Dresser C, Hartman A, Carroll M. Nutrient sources in the American diet: Quantitative data from the NHANES II survey. I. Vitamins and minerals. Am J Epidemiol 122:13–26. 1985.
94. Jacobs DR, Elmer PJ, Gorder DD, et al. Comparison of nutrient calculation systems. Am J Epidemiol 121:580–92. 1985.

Chapter 4 Dietary methodology issues related to energy balance measurement for NHANES III

by Dorothy Blair, Ph.D., Nutrition Program, College of Human Development, Pennsylvania State University

Introduction

The question most relevant to a discussion of dietary methodologies for measuring energy balance in the National Health and Nutrition Examination Survey (NHANES) III is not which method is most appropriate, but rather whether an attempt to measure energy balance is an appropriate aim for a large survey such as NHANES III. To answer this question we must first have a basic understanding of what energy balance is, what affects it, and who may be defined as being in energy balance. Then we should ask with which methods and under what conditions energy balance has been successfully measured and whether those methods and conditions can be duplicated or approximated in NHANES III.

This paper attempts to answer these questions and ends with a discussion of dietary methodologies that are most appropriate for the estimation of total energy intake for NHANES III.

Definition of energy balance

Energy balance can be defined as a close agreement between caloric intake and caloric expenditure. Garrow (1) defines energy balance as an agreement between energy intake and energy expenditure within ± 50 kilocalories per day (kcal/day). Although this is somewhat arbitrary, it demarcates the limits of our accuracy of measurement under very controlled circumstances. An imbalance of lesser magnitude would nevertheless have physiological consequences. A positive imbalance of 30 kcal/day could lead to an accumulation of 1 kilogram (kg) of body fat per year.

The test of energy balance in an individual is the maintenance of stable body fat stores. Fluctuations in carbohydrate stores may cause variations in body weight as great as 2 kg (1), with a 2,000 kcal fluctuation in body energy stores. Such changes in carbohydrate stores invalidate body weight as a measure of energy balance (2). The choice of stable body fat as the test for energy balance puts limits on who can be considered to be in energy balance. Growth- and reproductive-related changes in body fat stores rule out large segments of the population from the possibility of being in energy balance. The use of percent body fat is less valid in the context of energy balance than is total body fat in kilograms, as the percent body fat may fluctuate without a necessary change in total fat stores. An example is the aging-related loss of lean tissue, which leads to an increased percent of body tissue as fat without an increase in body energy stores.

Beaton makes a point worth noting: Energy balance is maintained in some cases at the expense of optimal physiological and mental functioning (3). An example is the adaptation of children to severely reduced energy intake through the cessation of growth and reduced activity. More commonly observed in affluent societies is the adaptation to low caloric intakes with severe and long-term dieting. Thus, energy balance is not of itself a desirable state but may be an adaptive response of the body to adverse circumstances.

Components and effectors of energy intake and expenditure

Table 1 shows the components of energy intake and energy expenditure and lists those factors that influence either side of the equation. Energy expenditure is under much tighter bounds than energy intake. It is a function of the fat-free mass (4,5), of short- and long-term caloric intake (6,7) and food composition, of illness and growth, and of movement (3). Movement would seem to be a discretionary category, but for most of us the caloric demands of the day are somewhat fixed. Leisuretime activity is the one dimension of energy expenditure we are free to define. Energy expenditure may vary from slightly greater than 1 × basal metabolic rate (BMR), assuming complete bed rest without food intake, to greater than 2 × BMR. Most of us live in the range of 1.3 to 1.5 × BMR (sedentary to light activity). Not even a 6-mile jog would hoist us into the heavy activity category. Mean caloric intake estimates of population groups that are under 1.4 × BMR must be viewed with suspicion.

Effectors of energy intake are much more varied and may be situational as well as under cognitive, emotional, or somatic controls. Physiological feedback mechanisms described by Van Itallie and Kissileff (8,9) and Booth (10) may be overridden by external stimuli—palatability, other people's behavior, and so forth. Conscious controls may be overridden by emotionality, alcohol, or drugs. To most, eating is a social, emotional, and sensual activity. It is synchronized with physiological needs only by dint of constant learned attention to physiological cues (10). For

Table 1. Components and effectors of energy balance

Intake effector	Caloric intake component	Body energy store	Caloric expenditure component	Expenditure effector
Conscious restraint or augmentation of intake	Fat	Body energy losses:	Basal needs	Fat-free mass
Palatability	Carbohydrate	Feces	Adaptive thermogenesis:	Long-term kilocalorie intake
Mood, stress, illness, drugs	Protein	Urine	Food induced	
Time and financial constraints	Alcohol		Nonspecific	Short- and long-term intake; meal composition
Weekend, weekday, season, holiday			Immune response, fever	
Somatic controls:			Growth, reproductive function; body building and training	Smoking, tension, drugs, thermal regulation
Energy expenditure and exercise				Cyclic variations
Physiological feedback			Movement and exercise	Fat-free mass; weight, day, season
Cyclical variations (menstrual cycle)				
Alcohol				

those who are less finely in tune with their body's needs, the salient cue may be the body composition change that is the negative outcome of ignoring earlier physiological cues.

The propensity toward obesity may be a function of the degree to which cognitive and somatic controls can be overpowered by external cues and emotional needs (11). The obese also appear to differ from "normal" persons in their exercise-induced eating. Increased exercise does not appear to influence the caloric intake of obese women, although normal persons increase intake in relation to exercise intensity (12). Intake may therefore be less tightly coupled to energy expenditure in obese than normal persons.

Most single effectors of both intake and energy expenditure have a coupled influence, so the chance for energy balance is heightened. For example, severe decrease in intake decreases both basal metabolism (6,7) and food-induced thermogenesis, which is approximately 10 percent of caloric intake (1) depending on meal composition. The energy expended in exercise drops as weight is lost (6). Smoking increases metabolism and increases consumption (13,14), although the increase in intake may be due to a relaxation of necessary restraint. The menstrual cycle causes a coupled variability in energy expenditure and intake (15). Exercise also increases intake over the long term in normal persons (12), though the relationship is not closely coupled within a 2-day period (16).

On the other hand, some effectors uncouple energy intake and expenditure. Illness increases metabolic requirements but reduces caloric intake through reduction in appetite (17). In some, stress has an appetite-depressing effect while muscular tension increases energy expenditure (18). A yearly cycle of weight loss and gain may occur because of seasonal and holiday overeating and exercise being reserved for the warmer months. Restrained eating and the conscious augmentation of energy expenditure could cause a particularly sharp energy imbalance.

Thus, although there are some physiological and cognitive reasons for a close correlation between energy intake and expenditure, there are other forces at work that disrupt this coupling and make a close match between energy intake and expenditure unlikely. It also seems to be true that the time required for energy balance to occur may differ between obese and lean because of a difference in intake cue responsiveness. In a sedentary society where food is plentiful, varied, and easily attained, energy balance on the short, measurable term may be the exception rather than the rule.

Precise estimations of energy balance and categorization of individuals

The point of measuring energy balance is not only to determine with some acceptable degree of reliability and validity those individuals who are in energy balance and those who are not. The broader goal would seem to be to increase our understanding of the mechanisms of body fat maintenance and the effects of body fat fluctuations on disease processes and outcomes. It must be stressed that energy balance, unless broadly defined and appropriately measured, is not the same as body fat maintenance.

It can be assumed that fat-stable individuals, whether normal or obese, fluctuate with some regularity around a set body composition. We do not know the periodicity of this fluctuation or the height of the arc, that is, the degree to which body composition change is tolerated by the individual before corrective mechanisms are employed. There are probably three basic types of body composition maintenance cycles, with many variations on these themes: (a) a close coupling of intake and output, resulting in slight weight fluctuations from day to day but no substantive change in body fat stores—in other words, true energy balance; (b) a short-term but measurable fluctuation of body fatness around a long-term stable body composition; and (c) a seasonal cycle of imbalance with an increase in winter holiday eating and cold-related decrease in exercise, followed by a reverse summer trend. These three styles are shown in figure 1. By definition, those individuals whose body composition fluctuates (types b and c) cannot be considered to be in energy balance, nor will they be measured as such by precise estimates of intake, output, or body composition assessments. Thus, those individuals whose body composition is stable by any other criteria will necessarily be considered out of energy balance if their body composition fluctuates regularly. The time period chosen for measurement will also have a profound effect on who is considered to be in energy balance. Those who

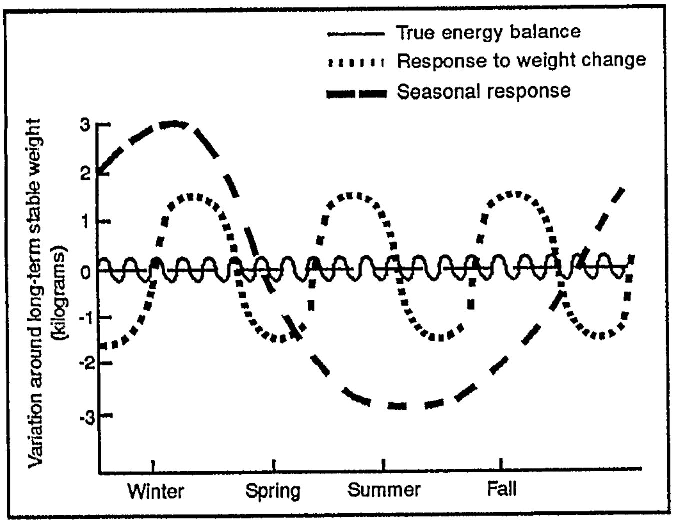


Figure 1. Three styles of long-term body composition maintenance

balance over a longer period and who are caught on the upswing or downswing will be considered out of balance.

Thus, many individuals who are fat stable in the long term will be designated as out of balance by precise, short-term studies with low tolerances for body composition change. The more stringent are the criteria for energy balance, the more false negatives will result. Positive and negative energy imbalances will exclude an individual from the maintenance criteria regardless of the outcome of that imbalance. Short-term energy balance studies have neither the sensitivity nor specificity to put individuals in categories that can serve a useful predictive function. The concept of body fat maintenance is more easily operationalized via serial measurements of body composition, which would be a more direct measure of the phenomenon of interest than energy intake and expenditure.

Attempts to measure energy balance: Realities of imprecision

The literature on energy balance can be divided into two categories: Those studies performed on small samples under fairly rigorous field conditions and those studies using methods more applicable to epidemiological studies such as NHANES. Metabolic chamber studies are not included in this review because of their impact on habitual behavior. The more rigorous field studies will be examined first to see if any techniques for measuring energy balance can yield results that are both valid and useful for the purposes of NHANES.

The most exacting field studies have obtained activity records and food records or weighed intakes over 7 days or longer. Energy expenditure has typically been calculated by the factorial method (19). Generally group means for energy intake have been very close to energy expenditure, but some individual subjects appear to be considerably out of balance over the timeframe of a week. Durnin and Brockway's review of six 7-day studies shows that, of 69 individuals studied, only 6 percent had daily intakes that correlated positively with daily energy expenditure (19). Twenty-six percent had significant differences in the mean energy intake minus energy expenditure after 7 days. Harries, Hobson, and Hollingsworth (20), reviewing studies done between 1955 and 1962, felt that 1 week was not a sufficient time to estimate energy balance because of the day-to-day variability in energy intake. The variability was considerably lower for energy expenditure than for intake.

None of the early estimates of energy balance used a body composition methodology more sophisticated than body weight. In an exceedingly long study of energy balance on a captive group of men in Antarctica, Acheson et al. (21) used a four-site skinfold estimate of body fat, repeated every week over a period of 6 to 12 months. However, none of their varied techniques for measuring energy balance was sufficiently accurate to predict individual fat gains or losses over the study period.

Both Acheson et al. (21) and Borel, Riley, and Snook (22) have utilized long-term caloric intake data along with long-term data on body composition to estimate the caloric intake required for energy balance. This method, the intake balance technique, corrects caloric intake for the caloric cost of changes in body composition. The technique has the advantage of correcting for body composition changes that may be the result of altered intake caused by the self-monitoring of food consumption. However, the correction factor is only as good as the method for estimating body composition. Certainly, skinfolds are too imprecise to be used in this context, but even densitometry could result in substantial errors in the estimation of the caloric cost of body composition change. The technique does have the advantage of eliminating the matching of two messy methodologies—dietary intake and energy expenditure—as a criterion for success.

Over a week, energy balance may be within ± 10–20 percent. The error of field energy expenditure techniques is at best ± 10 percent (1,23). Assuming accurate reporting, 3 weeks of dietary data via 24-hour recall or food record are necessary to estimate usual energy intakes with confidence limits of ± 10 percent (24). Poor subject cooperation and recall abilities, along with errors in food composition estimates and coding, can further decrease validity. In both cases, recording may change behavior. If randomness of error can be assumed, then use of group data would be possible, although the findings would be weak. If randomness cannot be assumed, then an error of this magnitude could obscure even a group comparison.

Thus, 7 days or more of continuous data collection of a type requiring intense subject cooperation and not a small intrusion on people's daily routine—methods totally unsuitable for NHANES—are not sufficiently valid to classify individuals into energy balance categories that are consistent with changes in their body fat stores.

The second area for examination in this review of the literature on energy balance studies is whether estimations of energy intake and expenditure more applicable to a large-scale epidemiological survey such as NHANES yield both valid and useful results. Because energy balance techniques are tedious for the subject and intrude on normal behavior, many investigators have looked for easier methods to measure one or both sides of the energy balance equation, with a resulting increase in the error of estimation. Usually the energy expenditure side of the equation suffers the most from approximation. The subject may spend tedious hours recording dietary intake over many days, although energy expenditure is estimated by use of an average daily heart rate and one heart rate-energy expenditure regression equation obtained in a laboratory under controlled conditions (6,25,26). More frequently, activity levels have become the surrogate measure for energy expenditure, presumably as the one variable in the energy expenditure equation under individual control. However, no attempt has been made to convert activity into calories expended, which leaves the impression that a given activity has an equal metabolic cost for all persons. The unconverted activity levels are then compared with caloric consumption.

A number of small- and large-scale studies have examined dietary intake and activity compared to body fatness. The dietary methods have ranged from 1 day or repeat 7-day food records and food histories to 24-hour recalls. Activity levels have been estimated through diaries, observation, or a mechanical device such as a pedometer. The results of these studies are as inconsistent as the methods are varied but can be grouped into three categories of results:

1. Body fat is related to decreased caloric intake and decreased activity—
   a. Johnson, Burke, and Mayer (1956): Diet histories, teenagers (27).
   b. Rose and Mayer (1968): Diet histories, infants (28).
   c. Hutson et al. (1965): 24-hour recall, adult males (29).
   d. Montoye et al. (1976): 24-hour recall, adult males (30).
   e. Baecke, van Staveren, and Burema (1983): 2-day food records, adult males (31).
2. Body fat is related to decreased caloric intake but not to differences in reported activity—
   a. Hutson et al. (1965): 24-hour recall, adult females (29).
   b. Baecke, van Staveren, and Burema (1983): 2-day food records, adult females (31).
   c. Huenemann (1967): 7-day food records, teenagers (32).
   d. Stefanik, Heald, and Mayer (1959): Diet histories, teenage boys (33).
   e. Bradfield, Paulos, and Grossman (1971): 3-day records, teenage girls (34).
3. Body fat is unrelated to caloric intake or activity—
   a. McCarthy (1966): 7-day diet record, adult females (35).
   b. Maxfield and Konishi (1966): 7-day diet record, adult females (36).

The methods used in these studies are so imprecise that biases in measurement between obese and lean subjects could easily be the major source of significant results, other errors being random. There was a tendency, at least in adults, to find no energy intake differences between obese and normal persons using a 7-day diet record, although 24-hour recalls routinely categorized the obese as eating less. This tendency was not true for obese children and adolescents, who were found to eat less than normal persons by every dietary method employed.

The temptation is strong to use very short-term dietary data from cross-sectional studies to draw conclusions relevant to energy balance and weight maintenance or even to categorize obese and lean individuals. We are ever hopeful that some relationship between caloric intake and body fat will be found, even if the meaning of those relationships is equivocal. The problems of drawing conclusions from such data can be illustrated by the analyses of the Tecumseh, Michigan, data by Montoye et al. (30) and of the NHANES I data by Braitman, Adlin, and Stanton (37). Both studies employed a 24-hour recall to approximate energy intake. The Tecumseh study employed a rigorous 30–60-minute recall of the preceding year's activities to classify 1,000 men into active, moderately active, or sedentary categories. NHANES I employed more modest estimates of perceived recreational and vocational activity on a scale from 0–2. These two activity scales were summed for an activity index ranging from 0 to 4.

Montoye et al. (30) found an inverse relationship between a year's average for work and leisure activity expressed as a work-to-basal ratio and a 1-day estimate of caloric intake. A diet history would have provided a more appropriate comparison to a year's worth of activity data. The authors state that their work-to-basal ratio is an estimate of energy expenditure, but in fact it tends to equalize the relationship between energy expenditure and activity between groups of different body compositions. The ratio penalizes the obese because the greater caloric cost of their exertion is balanced by their higher metabolic rate. It is not surprising that inactivity so defined was found to be related to a higher sum of skinfolds (table 2).

Table 2. Caloric intake and sum of 4 skinfolds by activity level in males: Tecumseh, Michigan

Activity category	*Kilocalories/ kilogram*	*N*	*Sum of skinfolds[1] in millimeters*	*N*
Most active	38.2	192	83.6	275
Intermediate	37.4	516	86.7	813
Least active	33.9	175	92.4	273

[1]Triceps, subscapular, suprailium, and juxtaumbilicus.

Reprinted with permission from *Journal of Chronic Disease,* 29, Montoye HJ, Block WD, Metzner HL, Keller JB, Habitual physical activity and serum lipids; Males age 16–64 in a total community, 1976, Pergamon Press, Elmsford, New York.

Caloric intake was standardized by kilograms body weight. A comparison based on caloric intake per fat-free mass would have kept the comparison congruous. An even more congruent comparison would be caloric intake compared with an estimate of average daily energy expenditure, based on an estimate of metabolism and the average basal-to-work ratio.

Montoye et al.'s sample ranged from 16 to 64 years of age. Decreasing activity and decreasing caloric intake are expected trends with aging, as is an increase in adipose tissue, at least until age 40, in men (38). Thus, the strength of these trends may be accounted for by aging itself and not any unique differences between obese and normal individuals. The lack of consistent significance across age groups strengthens this possibility.

The methodological difficulties, the standardization problems, and the comparison of a year's activity to 1 day of caloric intake all serve to make the data of Montoye et al. very difficult to interpret and accept at face value. Even if the comparisons were valid, could we infer from these data that increasing fatness is associated with decreasing caloric needs? No, because activity is only a fraction of energy needs, and because we know nothing about the variability in caloric intake in this sample. There is reason to believe that this variability is greater for the obese than the normal persons, as was stated earlier. The study tells us nothing at all about dietary variability and nothing about energy balance.

If more days of dietary intake were assessed or a diet history for the past year were performed, if the activity and metabolic data were converted to kilocalories, would the data then give us information on energy balance? The energy expenditure data might give an estimation of energy needs that, by one estimate, is accurate within ± 200 kcal (39). These are useful data. The estimate of caloric intake would be considerably more valid and useful if variability over time were considered (40). However the difference between the two estimates would not be of sufficient validity and reliability to determine the direction of imbalance, to predict weight maintenance, or to predict changes in body weight over time.

The data of Braitman, Adlin, and Stanton from NHANES I (37) are presented in table 3 for females and males. Reported caloric intake in females by 24-hour recall was inversely related to the percent optimum weight category. This relationship remained significant after adjustment for age and the perceived activity index. The relationship between caloric intake and weight category was not significant for men.

Can it be inferred from these data that obese females require fewer calories than normal females to maintain their weight? Is caloric intake in males unrelated to weight maintenance? The authors state that "there are two possible interpretations of these data: either obese adults eat no more than nonobese and therefore maintain their greater weight without excessive energy intake, or the estimates of food intake in this study differ in accuracy between obese and nonobese adults, either being underestimated by the former or overestimated by the latter." At least two other hypotheses are possible. One is that the obese are restrained eaters for the most part. Thus, a 24-hour recall may underrepresent the impact of their unrestrained days and underestimate true consumption. A second hypothesis is that the assumption of weight maintenance in each group is false. The tendency in cross-sectional studies is to assume that weight fluctuations in each group are equal, canceling each other, and that the caloric intake reported is for weight or energy store maintenance. There is cross-sectional evidence from the Health Examination Survey of 1960–62 and NHANES I, 1971–74, that over a 10-year time span, female age cohorts have increased energy stores in the form of body fat in all age groups but 55–64 years and 65–74 years. Men increased their subscapular skinfold in each age category, but triceps did not increase after age 35–44 years (41). Increases in internal body fat stores would be expected with increasing age. Other studies have shown significant changes in body composition over time (3). Obviously, 10-year changes are not in a league with the short-term fluctuations, which also could cause erroneous data interpretation, but both types of changes do lend credibility to the fourth hypothesis. We do not know which group, the obese or the normal, has a more consistent upward trend in body fat.

The degree of energy imbalance and the distribution of energy imbalance within a range of body fatness would be crucially important for the interpretation of the data of Braitman, Adlin, and Stanton. However, if longer term, valid data were available on energy intake and expenditure for NHANES III, we still could not determine energy balance with the degree of accuracy necessary to estimate the direction of true fat fluctuation. Only serial body composition estimates, the test of energy balance, would provide the information needed to establish both rapidly and unequivocally if a meaningful energy imbalance exists.

Table 3. Unadjusted and adjusted caloric intake at 5 levels of percent of optimum weight, by sex: United States, first National Health and Nutrition Examination Survey, 1971–75

Sex and percent of optimum weight	*Number of subjects*	*Unadjusted caloric intake*	*Caloric intake adjusted for physical activity and age* [1]
Female			
Less than 101	1,246	1,743	1,689
101–119	1,321	1,591	1,595
120–134	453	1,480	1,550
135–149	252	1,411	1,488
149 or more	245	1,478	1,525
Male			
Less than 101	708	2,423	2,359
101–119	1,241	2,366	2,375
120–134	512	2,264	2,310
135–149	155	2,264	2,310
149 or more	79	2,406	2,411

[1]For the analysis of covariance, for females, $F = 9.269$ and $p < 0.001$; for males, $F = 0.795$ and $p = 0.53 > 0.05$.

Appropriate caloric intake methodology for NHANES III

The precise and valid measurement of caloric consumption is an appropriate goal for NHANES III, but the limitations of available methods must restrain the uses to which the data are put. No method available is ideal and some are not even possible within the context of NHANES. Short- and long-term diaries and weighed records are not applicable because of their reliance on subject cooperation and recording abilities. They also tend to interfere with normal eating patterns. NHANES III must rely on methods that put the burden of work on the nutritionist rather than the interviewee. Rigorous quality control will have to be maintained in interviewer selection and training and through the collection of replicate data.

The following criteria should be considered when choosing the best dietary methodology for estimating energy consumption:

1. The data should reflect the full range of variability in the individual's diet, including weekdays, weekends, and other points of variability.

Dietary histories obtain a subject's estimate of habitual intake (42), although the method has also been used to obtain weekend and weekday estimates, which can then be weighted (43). The 24-hour recall has the advantage of giving data for discrete days so that group variability can be estimated if a suitable number of days are sampled (24,40).

NHANES II performed 24-hour recalls on Tuesdays through Saturdays (44), so the weekend variability in intake observed in other studies (45,46) was not recorded. Information on weekend-weekday patterns must be part of the dietary method. Each subject should have at least one 24-hour recall from a weekend day. If a weekend-weekday eating pattern is shown to exist, the intake estimate should be a weighted average of weekends and weekdays.

2. The method should minimize reporting bias, particularly differences between obese and normal subjects.

Twenty-four-hour recalls may underestimate caloric intake in obese versus normal subjects (47–49), but diet histories are subject to the same dependence on subject recall (50). When compared with a caloric intake that had maintained subject weight within ± 2 percent over at least 90 days, the diet history technique overestimated the normal subject's needs and underestimated the obese subject's needs (39). Whatever dietary method is employed, it should use special probing techniques aimed at uncovering the intake of high-calorie food items, such as snack foods, desserts, and alcohol, foods frequently underestimated (49,50). A food frequency cross-check designed to elucidate the consumption of these particular high-calorie food items may provide a useful probing tool.

3. The situation should maximize the subject's ability to remember both the types of foods consumed and the amounts.

Unless most meals are consumed away from the home, the home environment would provide the best site for the initial dietary interview. Here the setting and the contents of the refrigerator and cupboards provide memory cues. The size of portions can be approximated directly from the actual serving utensils. If interviews were computerized, information on usual container and plate sizes could be retained on the computer for subsequent interviews in the mobile examination centers and perhaps by telephone (51,52).

Because intakes on discrete days can help provide the estimates of day-to-day variability necessary to generate a statistical approximation of the distribution of usual intakes (24), repeat 24-hour recalls should be the method of choice for a cross-sectional comparison of group data. However, 2–3 days of dietary intake data cannot give a precise estimate of individual intake. As a result of the large intraindividual variation in caloric consumption, correlations and regressions of caloric intake with relevant physiological parameters are unlikely to yield more than marginally significant results (53). The probability of type two error—that is, the probability of accepting the null hypothesis when indeed the null hypothesis is false—is high.

In the right hands and with cooperative subjects, the diet history technique would seem to be a better choice than the 24-hour recall for prospective studies. Several authors have shown its repeatability over time, especially for energy intake (54–56). Changes in energy intake as estimated by diet history have been reasonably congruent with changes in body weight (57,58). A similar correlation using a 24-hour recall technique could not be expected. Diet histories give a higher estimate of caloric intake than 24-hour recalls (59,60), especially in the obese (48). One negative aspect of diet histories is the possible difficulty of obtaining information from school-age children (61). Diet histories are more time consuming to code and to convert to nutrient intake; therefore, they are more expensive to perform than 24-hour recalls. However, if an important part of an already expensive survey is to have comparable energy consumption data over time, the extra expense of thorough baseline dietary data would be compensated for by usable data.

Conclusion

NHANES can never hope to achieve the level of precision required for the measurement of energy intake and expenditure to estimate energy balance, nor is it clear that energy balance data would provide information more useful than body composition measured over time. The attempt to measure energy balance is an academic exercise more suited to testing the validity of methodologies than to meaningfully assessing individuals.

The accurate measurement of total energy intake is an appropriate and useful goal for NHANES III. Although no existing method is perfectly suited to the measurement of energy intake, NHANES could considerably improve the precision with which it presently measures this variable. Repeat 24-hour recalls with improved cross-checks are recommended for cross-sectional design. A diet history technique is recommended if a longitudinal design is adopted.

Acknowledgments

Thanks are due to Dr. Helen S. Wright for our conversations about dietary methodology and to Dr. John Beard for his comments on the manuscript.

References

1. Garrow JS. Energy balance and obesity in man. 2d ed. Amsterdam: Elsevier/North-Holland Biomedical Press. 1978.
2. Acheson KJ, Campbell IT, Edholm OG, et al. A longitudinal study of body weight and body fat changes in Antarctica. Am J Clin Nutr 33:972–7. 1980.
3. Beaton G. Energy in human nutrition: Perspectives and problems. Nutr Rev 41:325–40. 1983.
4. Cunningham JJ. A reanalysis of the factors influencing basal metabolic rate in normal adults. Am J Clin Nutr 33:2372–4. 1980.
5. Ravussin E, Burnand B, Schutz Y, et al. Twenty-four-hour energy expenditure and resting metabolic rate in obese, moderately obese, and control subjects. Am J Clin Nutr 35:566–73. 1982.
6. Warnold I, Carlgren G, Krotkiewski M. Energy expenditure and body composition during weight reduction in hyperplastic obese women. Am J Clin Nutr 31:750–63. 1978.
7. Liebel RL, Hirsch J. Diminished energy requirements in reduced-obese patients. Metabolism 33:164–70. 1984.
8. Van Itallie TB, Kissileff HR. The physiological control of energy intake: An econometric perspective. Am J Clin Nutr 38:978–88. 1983.
9. Van Itallie TB, Kissileff HR. Physiology of energy intake: An inventory control model. Am J Clin Nutr 42:914–23. 1985.
10. Booth DA. Acquired behavior controlling energy intake and output. In: Stunkard AJ, ed. Obesity. Philadelphia: WE Saunders. 1980.
11. Rodin J. The externality theory today. In: Stunkard AJ, ed. Obesity. Philadelphia: WE Saunders. 1980.
12. Pi-Sunyer TX, Woo R. Effect of exercise on food intake in human subjects. Am J Clin Nutr 42:983–90. 1985.
13. Hofstetter A, Schutz Y, Jequier E, et al. Increased 24-hour energy expenditure in cigarette smokers. N Engl J Med 314:79–82. 1986.
14. Jacobs DR Jr, Gottenborg S. Smoking and weight: The Minnesota Lipid Research Clinic. Am J Public Health 71:391–6. 1981.
15. Changes in nitrogen and energy metabolism during the menstrual cycle. Nutr Rev 41:116–8. 1983.
16. Durnin JVGA. "Appetite" and the relationships between expenditure and intake of calories in man. J Physiol 156:294–306. 1961.
17. Chandra RK. Nutrition, immunity, and infection: Present knowledge and future directions. Lancet 1:688–91. 1983.
18. Miller DS, Mumford P. Obesity: Physical activity and nutrition. Proc Nutr Soc 25:100–7. 1966.
19. Durnin JVGA, Brockway JM. Determination of the total daily energy expenditure in man by indirect calorimetry: Assessment of the accuracy of a modern technique. Brit J Nutr 13:41–53. 1959.
20. Harries JM, Hobson EA, Hollingsworth DF. Individual variations in energy expenditure and intake. Proc Nutr Soc 21:157–68. 1962.
21. Acheson KJ, Campbell IT, Edholm OG, et al. The measurement of daily energy expenditure—an evaluation of some techniques. Am J Clin Nutr 33:1155–64. 1980.
22. Borel NJ, Riley RE, Snook JT. Estimation of energy expenditure and maintenance energy requirements of college-age men and women. Am J Clin Nutr 40:1264–72. 1984.
23. Passmore R. Energy balances in man. Proc Nutr Soc 26:97–101. 1967.
24. National Academy of Sciences, National Research Council, Subcommittee on Criteria for Dietary Evaluation, Coordinating Committee on Evaluation of Food Consumption Surveys. Nutrient adequacy: Assessment using food consumption surveys. Washington: National Academy Press. 1986.
25. Bradfield RB, Jorden M. Energy expenditure of obese women during weight loss. Am J Clin Nutr 25:971–5. 1972.
26. Griffiths M, Payne PR. Energy expenditure in small children of obese and non-obese parents. Nature 260:698–700. 1976.
27. Johnson ML, Burke S, Mayer J. Relative importance of inactivity and overeating in the energy balance of obese high school girls. Am J Clin Nutr 4:37–44. 1956.
28. Rose HE, Mayer J. Activity, caloric intake, fat storage, and energy balance of infants. Pediatrics 41:18–28. 1968.
29. Hutson EM, Cohen NL, Kunkel ND, et al. Measures of body fat and related factors in normal adults. J Am Diet Assoc 47:179–86. 1965.
30. Montoye HJ, Block WD, Metzner HL, Keller JB. Habitual physical activity and serum lipids: Males age 16–64 in a total community. J Chron Dis 29:697–709. 1976.
31. Baecke JAH, van Staveren WA, Burema J. Food consumption, habitual physical activity, and body fatness in young Dutch adults. Am J Clin Nutr 37:278–86. 1983.
32. Huenemann RL. Teen-agers' activities and attitudes toward activity. J Am Diet Assoc 51:433–40. 1967.
33. Stefanik PA, Heald FP, Mayer J. Caloric intake in relation to energy output in obese and nonobese adolescent boys. Am J Clin Nutr 7:55–61. 1959.
34. Bradfield RB, Paulos J, Grossman L. Energy expenditure and heart rate of obese high school girls. Am J Clin Nutr 24:1482–8. 1971.
35. McCarthy NC. Dietary and activity patterns of obese women in Trinidad. J Am Diet Assoc 48:33–7. 1966.
36. Maxfield E, Konishi F. Patterns of food intake and physical activity in obesity. J Am Diet Assoc 49:406–8. 1966.
37. Braitman LE, Adlin EV, Stanton JL Jr. Obesity and caloric intake: The National Health and Nutrition Examination Survey of 1971–1975 (HANES I). J Chron Dis 9:727–32. 1985.
38. Blair D, Habicht JP, Sims EAH, et al. Evidence for an increased risk for hypertension with centrally located body fat and the effect of race and sex on this risk. Am J Epidemiol 119:526–40. 1984.

39. Mahalko JR, Johnson LK. Accuracy of predictions of long-term energy needs. J Am Diet Assoc 77:557–61. 1980.
40. Beaton GH, Milner BA, Corey P, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J Clin Nutr 32:2546–59. 1979.
41. Bowen PE, Custer PB. Reference values and age related trends for arm muscle area, arm fat area and sum of skinfolds for United States adults. J Am Coll Nutr 3:357–76. 1984.
42. Beal VA. The nutritional history in longitudinal research. J Am Diet Assoc 51:426–32. 1967.
43. van Staveren WA, deBoer JO, Burema J. Validity and reproducibility of a dietary history method estimating the usual food intake during one month. Am J Clin Nutr 42:554–9. 1985.
44. McDowell A, Engel A, Massey JT, Maurer K. Plan and operation of the second National Health and Nutrition Examination Survey, 1976–80. National Center for Health Statistics. Vital Health Stat 1(15). 1981.
45. Beaton GH, Milner J, McGuire V, et al. Source of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Carbohydrate sources, vitamins and minerals. Am J Clin Nutr 37:986–95. 1983.
46. Richard L, Roberge AG. Comparison of caloric and nutrient intake of adults during week and weekend days. Nutr Res 2:661–8. 1982.
47. Beaudoin R, Mayer J. Food intakes of obese and non-obese women. J Am Diet Assoc 29:29–33. 1953.
48. van den Berg AS, Mayer J. Comparison of a one-day food record and research dietary history on a group of obese pregnant women. J Am Diet Assoc 30:1239–44. 1954.
49. Linusson EEI, Sanjur D, Erickson EC. Validating the 24-hour recall method as a dietary survey tool. Arch Latinoam Nutr 24:277–94. 1975.
50. Bray GA, Zachary B, Dahms WT, et al. Eating patterns of massively obese individuals: Direct vs. indirect measurements. J Am Diet Assoc 72:24–7. 1978.
51. Krantzler NJ, Mullen BJ, Schutz HG, et al. Validity of telephoned diet recalls and records for assessment of individual food intake. Am J Clin Nutr 36:1234–42. 1982.
52. Posner BM, Borman CL, Morgan JL, et al. The validity of a telephone administered 24-hour dietary recall methodology. Am J Clin Nutr 36:546–53. 1982.
53. Sempos CT, Johnson NE, Smith EL, Gilligan C. Effects of intraindividual and interindividual variation in repeated dietary records. Am J Epidemiol 121:120–30. 1985.
54. Dawber TR, Pearson G, Anderson P, et al. Dietary assessment in the epidemiologic study of coronary heart disease: The Framingham study. II. Reliability of measurement. Am J Clin Nutr 11:226–34. 1962.
55. Reshef A, Epstein LM. Reliability of a dietary questionnaire. Am J Clin Nutr 25:91–5. 1972.
56. Jain M, Howe GR, Johnson KC, Miller AB. Evaluation of a diet history questionnaire for epidemiologic studies. Am J Epidemiol 111:212–9. 1980.
57. Trulson NF, McCann NB. Comparison of dietary survey methods. J Am Diet Assoc 35:672–6. 1959.
58. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 115:492–505. 1982.
59. Young CM, Hagan GC, Tucker RE, Foster WD. A comparison of dietary study methods. II. Dietary history vs. seven-day record vs. 24-hour recall. J Am Diet Assoc 28:218–21. 1952.
60. Morgan RW, Jain M, Miller AB, et al. A comparison of dietary methods in epidemiological studies. Am J Epidemiol 107:488–98. 1978.
61. Rasanen L. Nutrition survey of Finnish rural children. VI. Methodological study comparing the 24-hour recall and the dietary history interview. Am J Clin Nutr 32:2560–7. 1979.

Chapter 5
Critique of studies of the relationship between diet and osteoporosis

by Nancy E. Johnson, Ph.D., Department of Food Science and Human Nutrition, University of Hawaii

Introduction

Osteoporosis is an age-related disorder characterized by decreased bone mass and by increased susceptibility to fractures, especially fractures of the vertebrae, distal radius, and proximal femur (1,2). Heaney (3) points out that fracture and decreased skeletal mass are distinct and ought not to be confused. The relationship between the two is still inadequately understood. For most physicians, fracture itself is the principal diagnostic evidence of osteoporosis, with about 1.3 million fractures attributable to this condition. The risk of osteoporosis developing increases with age, especially after 45 years, and is higher in women than in men and in white than in black persons. The frequency with which age-related fractures occur in different geographic areas varies enormously. Deaths from falls are reportedly less common in areas with naturally fluoridated water than in other regions, but the mortality data are unreliable. Incidence rates of fractures seem to be higher among white than among other persons regardless of the geographic area involved (4). The true reasons for racial differences are unknown, although differences in bone density have been described, with American black persons said to have substantially greater levels of bone density than white persons of the same age and sex.

Melton and Riggs (4) point out that, in general, people residing in geographic areas where diets are deficient in dietary calcium, protein, and vitamin D have lower fracture rates, whereas the incidence is highest in developed countries where diets are better. Thus, the observation of lower fracture rates in developing countries is the opposite of what one would expect. Of course, people residing in developing countries have lower lifespans than those in most developed countries, where environmental and health conditions are different.

Clinical studies of patients with fractures of the hip, distal radius, proximal humerus, and vertebra have generally found some evidence of diminished bone density at the site of the fracture. It has been shown that fracture rates are greater for women, who have lower bone density levels than men, and that the rates increase with advancing age among both sexes as bone density diminishes (4). Thus, the risk of fracture and osteoporosis increases as bone density declines. In a geographically defined population of women ages 55–80 years from two small demographically similar communities in the United States, the physician-documented fracture rate was 27 percent in women ages 55–64 years and 39 percent in women ages 65–80 years. The mean bone density of women with fractures was significantly lower than that of women with no fractures. Women from the community with low-calcium water reported more vertebral fractures than women from the high-calcium community (5).

It has been suggested that the different fracture syndromes generally associated with osteoporosis might represent separate pathogenetic mechanisms. Crush fractures in postmenopausal women represent early trabecular bone loss; hip fractures represent the loss of both cortical and trabecular bone from the proximal femur. The data obtained from studying patterns of bone loss from one part of the skeleton may not apply to all fracture syndromes of interest. Parfitt (6) proposes two structurally different forms of bone loss. Rapid bone loss leads to changes in structural elements of trabecular bone, an increase in the size of marrow cavities, and discontinuity in bone structure. Changes also occur in cortical bone. These structural characteristics reduce the strength of the bones to a greater extent than the reduction in the amount of bone by itself would suggest. Slow bone loss results from incomplete refilling by osteoblasts of resorption cavities of normal or reduced size. This leads to simple thinning of structural elements in both trabecular and cortical bone and reduces the strength of the bone in proportion to the reduction in the amount of bone. Ruegsegger et al. (7) found that postmenopausal osteoporosis appears to develop in phases of relative stability and of acute bone loss. There is a steplike pattern of trabecular bone loss. Also, the yearly trabecular bone loss was considerably higher in osteoporotic patients (2.7 percent) than in healthy postmenopausal women (0.95 percent). These findings suggest that a rate of loss of trabecular bone that exceeds normal age-related bone loss can lead to osteoporosis.

Information on bone mineral content is important for monitoring age-related bone loss. Measurements may be performed to assess cortical or trabecular bone mineral loss resulting from accelerated bone resorption or decreased bone formation to provide a quantitative result that can be used as a predictor of fracture risk. There are differences in bone loss patterns in different bones, and bone loss occurs with different rates on different bone surfaces even within the same bone. Furthermore, these patterns vary in prominence in different individuals (8).

Some areas of compact bone, particularly the metacarpals, are poorly correlated with the spine and areas of trabecular bone. Other areas of compact bone, the radius, for example, provide reasonable prediction of femoral neck mass. However, with aging and with bone disease, the degree of association among these areas decreases and the errors of prediction increase. This is the result of preferential loss of trabecular bone in the axial skeleton (9). However, when low bone mass is found at any measured site, there probably has been a general loss of bone (9,10). According to Heaney (3), the notion of osteoporosis implies a causal relationship between decreased bone mass and increased fragility. It is now clear that decreased mass is important, and probably the most important factor in osteoporotic fracture. Bone mass is a good discriminator between those who are likely to fracture and those who are not.

Lifestyle factors such as cigarette smoking, high caffeine intakes, and alcoholism play a role in aggravating the tendency to age-related bone loss. Other factors include the following (3):

- Slight or slender build.
- Fair skin.
- Family history of osteoporosis or osteoporotic fracture.
- Small muscle mass.
- Sedentary lifestyle.
- Small peak adult bone mass at age 35 years
- Low calcium intake.
- Early menopause or oophorectomy.
- One or more prior osteoporotic fractures.

It has been suggested that those who have achieved a larger bone mass by the fourth decade are slower to evidence clinical expression of bone loss later in life (11). Thus, the absolute amount of bone mineral present in the skeleton may dictate the subsequent development of osteoporosis and fractures.

Dietary factors and bone

There are a number of reviews that provide general information about nutrients and hormonal influences on bone (2,3,12–17). The relationship between dietary factors and bone mineral content or osteoporosis has not been established unequivocally. Research on the role of dietary factors in the causation of bone loss leading to osteoporosis has focused primarily on calcium, although a few other nutrients have been examined for their role in affecting calcium balance or calcium metabolism. The nutrients that generally have been associated with bone include calcium, vitamin D, phosphorus, protein, sodium, fluoride, vitamin C, vitamin A, magnesium, zinc, and total calorie intake. Other nonnutrient components of diet may also affect bone metabolism.

In an epidemiological study of Japanese-American men and women living in Hawaii, 1,208 men and 912 women were examined (18). Both 24-hour dietary recalls and the frequency of eating selected foods containing large amounts of calcium during a previous week were collected. An analysis of these cross-sectional data indicated that dietary or supplemental intake of milk, calcium, magnesium, vitamin A, vitamin C, vitamin D, and caffeine were significantly related to bone mineral content at one or more skeletal sites, even after adjusting for confounding factors such as age, weight, height, exercise, history of nonviolent fractures, use of thiazides, and use of estrogen. However, the effects of diet were modest in comparison to effects of the other factors. For example, dietary calcium intake could explain only 0.5 percent of the variation in the bone mineral content of the distal radius in men, and all dietary variables contributed R^2 values of 0.3–0.8 percent. In this study, the sodium-to-calcium ratio was significantly and negatively correlated with bone mineral content at all skeletal sites only in men. Caffeine intake was negatively correlated with the radius and ulna only in women. There was no significant association of supplemental calcium intake with bone mineral content. The other interesting point of note was that in women there was no relationship between calcium from the 24-hour dietary recall and any of the bone measurements, although there was a significant relationship with milk, and also with weekly calcium intake. For men there was a significant relationship between bone measurements and both 24-hour dietary calcium and weekly milk or calcium intake. However, in both groups there was no correlation between any of the bone measurements and either supplemental calcium or dietary plus supplemental calcium.

Similarly, Sowers, Wallace, and Lemke, who examined midradius bone density in a geographically defined population of 325 women living in Iowa, did not find a correlation between total calcium intake and bone density (5). However, mean bone density was greater in persons whose calcium intake was greater than 800 milligrams per day (mg/day) when this intake was consumed concurrently with vitamin D in amounts greater than 400 international units. The vitamin D intake was significantly correlated with bone density in this study, but supplemental vitamin D was negatively correlated with all bones measured in the Hawaii study (18). There was no significant difference in bone density between women from the high-calcium community and those from the lower calcium community in the Iowa study (5). The level of calcium in the water in one community was such that long-term daily intakes could be at a level of 1,200 to 1,500 mg/day. The total average calcium intake in the low-calcium area was 745 mg of calcium for food with a total intake of 964 mg; in high-calcium areas the average from food was 740 mg, and the high-calcium area averaged 1,329 mg when water was included. As these average intake levels are relatively high, it is not really surprising that there is not much difference between communities because women in both communities had intake levels well above both the RDA (Recommended Dietary Allowance) figure and some estimates that have been made of average calcium requirements.

Freudenheim, Johnson, and Smith recently completed a 4-year study of bone loss in women 35–65 years of age (19). Diet was monitored for three of the four years, and usual intakes of individuals were determined. When a cross-sectional analysis of the data was carried out, the only nutrients related to bone mineral content at the beginning of the study were vitamin C and niacin. Examination of bone loss in these women over the 4-year period revealed that calcium, protein, phosphorus, zinc, folate, and magnesium were associated with less loss. In a few cases excessive vitamin C was associated with increased loss.

Dietary methods related to study of osteoporosis

The osteoporosis literature contains a variety of dietary methods that have been used to associate nutrient intakes with bone mineral measurements or with incidence of fracture. These methods include single 24-hour recalls, 3-day records, diet histories, quantitative frequencies for a week, retrospective frequency at various stages of life, and serial records. The methods used are sometimes poorly described. Some of the methods provide an estimate of usual dietary intakes; the 24-hour recalls and 3-day records provide an estimate of intake at that particular time and are likely not a good estimate of usual intake, given the large amount of intraindividual variance in dietary surveys (20,21). Studies that have been done are primarily of adults, often aging adults, and rarely of infants, children, and adolescents.

The ability of different dietary methods to provide important information about relationships between diet and bone depends upon the kind of information sought about both diet and bone, the statistical methods used to analyze the data, and the nutrient under study. Retrospective dietary methods that attempt to identify intakes during childhood, adolescence, and early adulthood need to reflect usual intake over a specific time period. These might be the most appropriate methods for cross-sectional studies and as part of case-control studies. For a cohort study, in which a series of measurements are made on relevant bones, if the information desired is rate of bone loss or changes in bone density during the time of the study, a different dietary methodology would be desirable. It would be important to determine the usual diet during the period of bone changes. This can be accomplished by monitoring the diet and other related constituents through a series of observations. The type of dietary method used needs to be one that describes usual quantitative intake with reasonable accuracy. For comparative purposes and also because one method provides slightly different results from another method, the same method should be used throughout the study, and some baseline observations should be made using the identical method.

A single observation on diet might not provide a good correlation between diet and bone because of the high intraindividual variance of dietary data that would bias the coefficient (r) toward zero. Intraindividual variance is different for different nutrients, so the number of days, if single-day records or recalls are used, will be different if usual diet is estimated (20,21).

Differences between frequency and 24-hour recall methods can lead to different conclusions. In a recent study (18) it could be concluded that the frequency method for estimating calcium obtained through major food sources, primarily milk or dairy products, was best because it represented usual milk intake, and likely habitual milk intake that may have gone back to early adulthood or adolescent food practices. This is illustrated in data presented by Yano et al. (18) using calcium as an example. For 912 Hawaiian women with lower bone mineral content than other comparable groups in the United States, daily calcium intake estimated by the 24-hour recall was not correlated with radius, ulna, or os calcis. Morever, neither daily calcium supplement intake nor the sum of supplement and 24-hour dietary intake was correlated with any bone measurement. However, daily milk intake was correlated with the distal radius, and weekly calcium intake, which was estimated using frequency of consumption of standard portion sizes of foods containing large amounts of calcium, was significantly correlated with the distal radius. Because we have developed the habit of equating calcium with milk and vice versa, we would conclude that the frequency method is better because it measures usual intake of calcium and thereby eliminates the problem of intraindividual variance biasing the correlation coefficient when only a single observation, the 24-hour recall, is made. Therefore, in spite of the fact that daily calcium intake, supplemental calcium intake, and daily plus supplemental calcium intake are not correlated, we might conclude that we need to use a frequency method to obtain better estimates of calcium consumption.

There are other possibilities. It is possible that the correlation between milk and the distal radius is the critical factor. It is also possible that it is not calcium alone that is important but either another nutrient found in milk or a combination of nutrients that enables calcium to be absorbed more efficiently or bone mineral formation to take place more rapidly. In our own studies (Freudenheim, Johnson, and Smith) of women 35–65 years of age whom we followed for a 4-year period, we found that less bone loss in the radius and improvement in loss was significantly correlated with phosphorus, magnesium, and protein, not calcium, in a small number of premenopausal women and with energy, protein, calcium, phosphorus, and zinc in 33 postmenopausal women (19). When a 1.5-gram calcium supplement was added to the usual diet of half the women, the relationships between diet and radius bone mineral loss were erased and not even calcium was significantly correlated with bone. The mixture of nutrients seemed to be important in preventing bone loss. Our dietary method was to obtain two precoded, prevalidated food records at two times during each month for a 3-year period. Each of the 2 days was selected using a table of random numbers; 29 days later the next record was recorded. Thus, serial records were kept by the women and usual diet was monitored for three of the four years.

Some of these points can be illustrated further with another widely quoted study that was conducted in Yugoslavia (22). The diet methodology was called a diet history about usual patterns of eating, with questions on different food items. One person per household was interviewed, the information was extrapolated to the rest of the household, and identical household measures and food portions were used for the questionnaire. Thus, one assumes that a quantified history was obtained. Hand radiographs were used to estimate bone mass; femur and forearm fractures were recorded over a 6-year period. Two different districts, one low calcium and one high calcium, were chosen. In the high-calcium district, calcium and phosphorus were both two times as high, protein was about 50 percent higher, and kilocalories were 10 percent higher than in the low-calcium district. The fracture rate was higher in the low-calcium, low-phosphorus district, with good correlation between bone density and fracture. The authors concluded that the effect was a calcium effect. Because of high intercorrelations between calcium and phosphorus ($r = 0.961$ in the Freudenheim, Johnson, and Smith study (19)), the effective nutrient could have been phosphorus, an interaction between both nutrients, or a mixture of two or more nutrients.

This raises the question of multicollinearity between nutrients and between calories and almost all nutrients. In the Freudenheim, Johnson, and Smith study (19), the correlation between protein and kilocalories was 0.85; phosphorus and protein, 0.91; and calcium and phosphorus, 0.96. With intercorrelations between nutrients, it is almost impossible to select one nutrient as the most important one. In a study of diet and blood pressure, Reed et al. (23) found that, although the association with potassium intake was relatively stronger than associations with other nutrients, the intake of potassium was so highly correlated with intakes of calcium, milk, and protein that it was not statistically possible to identify the independent association of potassium and blood pressure. Although they could not easily separate the diet items, there was some support for the idea that the mixture of nutrients was important. The concept that a food item or specific mixture of foods is associated with retention of bone mass needs to be kept in mind for the third National Health and Nutrition Examination Survey (NHANES III).

There is another problem that needs to be considered: That is, the present bone mass may be determined by early dietary habits and the total amount of bone laid down before adult bone loss begins its relentless course. Therefore, current habits may not be correlated with current bone mass.

Different dietary methodologies may be necessary to estimate the different nutrients or foods related to the development of osteoporosis. The usual intake of nutrients, such as vitamins A and C, with high coefficients of variation, much of which can be due to high intraindividual variance, may be better estimated using quantified frequency methods. However, estimates of energy intakes and the many nutrients highly correlated with energy and present in small amounts in a wide variety of foods may be better estimated by using a more precise method and by monitoring the quantitative consumption on a regular basis. Thus, for associating bone loss over time, which will require a series of bone density measurements to determine rates of loss for individuals, quantitative nutrient intakes over the same time span will be needed. A combination of serial observations and frequency of consumption of certain food items that are good sources of vitamins A, C, and D and fluoride may be the best procedure.

For some dietary components such as fluoride, the analytic procedures are very poor for foods, although procedures are adequate for water supplies and other solutions. There are other mineral contributions to water supplies and these contributions can be substantial. Estimates of usual water intake would be useful. Also, fluid intakes should be estimated, as tea is a good source of fluoride; coffee provides caffeine, which increases excretion of calcium; and soda and other beverages may provide phosphorus and other compounds that may be found to have an effect on bone.

Dietary vitamin D can be estimated by using a few foods, but amount of sunlight and supplemental intake will have an impact on the amount of circulating $25(OH)D_3$ and $1,25(OH)_2D_3$. The $1,25(OH)_2D_3$ circulating in the blood will have the greatest impact on calcium absorption, bone formation, and bone resorption.

The factor of exercise and its relationship to bone mineral content needs to be considered when data are analyzed. Supplemental estrogen therapy, thiazide ingestion, and other factors need to be considered along with the dietary assessment. Many women are now taking various forms of calcium supplements, and a high percentage of people (close to 40 percent) use vitamin or mineral supplements.

The question of applicability of different dietary methods to various study designs, such as case-control, cross-sectional, and cohort, also must include consideration of the bone measurement observations—that is, is the intent to examine bone mineral content at one point in time or to determine the rate of bone loss? The bone mineral content measured at a single point in time may be determined by early dietary patterns, genetic factors, previous growth retardation periods, and other nonnutritional considerations. Few relationships have been found between current diet and cross-sectional measurements of bone. In the larger population group studied in Hawaii (18) the nutrients related to bone mineral content in cross-sectional analysis were those found in milk, with the exception of vitamin C in women. In Iowa (5) only vitamins D and C were positively related to bone, and in Wisconsin (19) only vitamin C and niacin were related.

For cohort studies the loss of bone mineral over time (or maintenance of bone density in adults) and formation of bone in children may be the best measure of potential for the development of osteoporosis. For these studies, monitoring both nutrient and energy intakes over time would be desirable. Morgan et al. (24), using survey data that included 12 daily records per individual from the Exploratory Study of Longitudinal Measures of Individual Food Intake of the U.S. Department of Agriculture (USDA), found that for food energy estimates and for "small eaters" or low consumers, little additional information is obtained after 6 days. For 14 larger eaters, the greatest gains in accuracy occurred prior to day 7, but additional observations did improve the estimates. Sempos et al. (20) found that 7 repeated records are needed for estimating the true population correlation $E\rho$, and 15 are needed for the slope ($E\beta$). For calcium, phosphorus, and magnesium, 6 days are needed to estimate $E\rho$. Vitamins A and C are more difficult to estimate, as the ratios of intraindividual variance to interindividual variance are higher (20).

The development and validation of cost-effective methods to monitor nutrient and energy intakes are necessary prerequisites to any large-scale study. Morgan et al. (24) also examined nine methods of assessing dietary intake status. They found it necessary to have an interviewer for a 1-day recall with either a return to pick up a 2-day diary or telephone recalls after the initial contact. A mail method provided response rates of less than 50 percent. Personal interviews provided the most consistent reporting of mean numbers of total food items. Sempos et al. (20) used an initial instruction method, which was ultimately taped, and mail return of the structured forms, with postcard and phone reminders. The return rate was 80 percent in a highly motivated group of women. A first-day interview and recall by telephone provided the most consistent reporting of food energy intake according to Morgan et al. (24). Telephone interviews provided the best food energy intake data with the most reasonable effort.

Methods that work for total calories will work for nutrients that are highly correlated with total calories. For those nutrients not associated so directly with calories, such as vitamins A and C, frequency methods need to be used.

As noted previously, it may be important to associate not only nutrients but also a food or a group of foods with the bone measurements. Recent studies have focused on examining different ways to assess not only nutrient intakes but food patterns and food item and food group intakes. Shortened intake estimation procedures have been developed for the purpose of determining the relative consumption of dietary components. These procedures make use of a limited data base and are based on the concepts that a few food items that are frequently used by a population constitute "core" foods (25,26) and can be used to assess nutrient intakes or food intakes. For most of the frequency studies a short list of foods or food groups can be developed. Willett et al. (27) used 99 foods, Johnson et al. have used short forms of from 68 to 183 foods (25), and Byers et al. (26) used 128 foods in an abbreviated interview form. Variability in nutrient intake can be explained by a small number of foods in some cases (26).

The ability of a nonquantitative frequency to substitute for quantitative methods is limited. In one study (28) there was reasonable agreement for food items, but intakes of food groups and nutrients did not provide good relative agreement, and group absolute agreement was found even less frequently. For individuals there was no absolute agreement for nutrient intakes, and agreement for food groups and food items was rare. Frequency data could be used to rank groups of individuals according to mean intake of food items but not of food groups or nutrients. Where effects are large, group dietary methods can tolerate a larger error component. However, when dietary and physiological relationships are weakly associated, as they have been in almost every study of diet and disease, the dietary estimate must accurately reflect an individual's intake. Individual contribution of dietary variables to multiple regression models of diet and bone mineral content can be small, from 0.3 to 0.8 percent (18).

For cross-sectional studies that include measures of current bone mineral content, some estimate of usual diet could be important. With studies of bone, these data may need to be analyzed in quantiles and using group means. Frequency methods involving specified foods and portion sizes have been used by several investigators (27,29). However, comparisons of these methods with those obtained using more precise methods have yielded variable results. Early studies by Johnson et al. (25,30) showed some overestimation using frequency and some but not all good food group comparisons. Thus, a single observation, or at least three measurement periods using a method appropriate to other studies, may be desirable.

Osteoporosis occurs when there is either rapid bone loss during the adult and aging years or slow bone loss during an extended lifespan. Although previous diet may dictate total skeletal mass and bone mineral content, if one could maintain that skeletal mass, one would be unlikely to develop osteoporosis. Thus, it would seem reasonable in case-control and cohort studies to examine rates of bone loss over time. Because the loss rates are usually small, a study interval of from 3 to 5 years is needed to determine loss rates.

For monitoring diet during periods of loss, single observations, whether they are recall or structured records, could be used. Morgan et al. (24) and others have shown that telephone interviews are feasible, and this may be the best way to collect these data for most of the nutrients, and especially for calories and those nutrients highly correlated with calories.

Since 1974 Johnson, Nitzke, and VandeBerg have used precoded, structured dietary record forms that provided good estimates for the specific nutrients under study (30). The forms were based on frequency of consumption of food items by the population group being studied. For a study of nursing homes (25), Johnson et al. used the cycle menus to develop a data base of 288 foods most frequently served to residents and validated the form against weighed food intake measurements. We found the forms reasonably accurate on a percent paired-difference basis, although there were a number of statistically significant differences. The advantage of using these forms is that they can be processed rapidly and inexpensively. They do not serve to provide good estimates of all nutrients, but for calcium and protein there was no significant difference between weighed intake and the form; and for calories, vitamin A, vitamin C, magnesium, and zinc, differences were less than 5 percent. To include foods consumed but not on the form, we have a catchall "other foods" category where any item consumed can be written on the form.

Recommendations

It is likely that information about current diet will not provide information relevant to bone mineral content as measured at the same time, with a few exceptions. Consumption of any item, food, beverage, or vitamin-mineral supplement currently consumed on a regular basis could affect bone mineral content. However, this information can be obtained by asking about a usual diet with frequency-of-consumption questions. Contributions of minerals from current and past water supplies need to be determined.

Early dietary patterns, particularly those related to milk, dairy products, and other calcium-contributing foods, may be useful in association with current bone mineral content.

In order to make NHANES III data comparable with past NHANES and USDA data, at least some common measurement method should be used. Because the 24-hour recall method has been the one consistently used, that could be considered. I prefer a day's record to minimize forgetting errors, and this has also been used in the USDA surveys, but not to my knowledge in NHANES.

Although 24-hour recalls on single-day records could be used to classify individuals into quantiles, estimates of the classification errors involved need to be made. Attempts to correlate single measurements with bone mineral content have not proved very fruitful.

Because dietary and bone mineral relationships have not been clearly and conclusively shown with cross-sectional analysis and because the rate of bone loss is as important as total bone mass, especially for women, it is important to monitor changes in both bone mineral content and diet. Therefore, a series of observations on an individual should be made over the same time period that bone is being monitored. For this, either a 24-hour recall or record observation for a series of randomly selected days is recommended. From the studies that have been conducted, it seems that 6 days would be best for best estimates. The best response seems to be obtained with an original contact and telephone followups, probably with the same interviewer contacting an individual to take advantage of the rapport established during the first interview. Of the six recalls or records, data from the first 3 days can be used to classify individuals into quantiles with more accuracy than a single record or recall.

Precoded, shortened food record forms have been found to facilitate data processing, decrease processing costs, and increase the rapidity with which statistical analysis of the nutritional status indicators and diets can be accomplished. With computer equipment currently available, field data could be instantly input to a central file, and the central file could be immediately converted to food group and nutrient data, with the ultimate form of the data in measures that are immediately amenable to statistical procedures.

In summary, methods not too different from those previously used would be the best considering all constraints. In addition, serial measurements should be made to monitor intake. If new frequency methods are used, especially shortened forms based on fewer foods or nutrients, these methods and forms should be pretested with standard recording and recall methods to determine where the errors are likely to be and how great the differences are likely to be.

1. Information about specific food items, previous dietary patterns, and composition of the water supply could occur at the first interview. As there is a need to collect information about the history of vitamins A and C intakes for cancer risk factors, this information could be used for the osteoporosis component as well. Retrospective questions about dairy food consumption may provide information that can be associated with bone mineral content.
2. To develop baseline data for monitoring, a series of random observations would need to be made about the time of the initial measurement. The exact number of observations depends on the error or discrepancy between those observations and an estimate of the usual intake. The methodology for comparative purposes must be consistent. The "apples and oranges" effect manifests itself when two different methods are used within a single study.
3. The quantitative method selected for baseline measurement and for subsequent observations should be validated by testing the instrument against weighed food intakes and by chemical analysis of the weighed foods. This does not deal with the bioavailability problem in diets, but it does provide an idea of how accurately the methods portray the true intake of the specified nutrients. Accuracy of reporting and using the questionnaire should also be tested in a study population.
4. For classification of individuals into quantiles, at least three observations (and at least six observations for nutrients with relatively small variance) are recommended to obtain reasonable estimates of intakes over time. For high-variance components of the diet, quantified frequency questions may be the only way to obtain reasonable estimates of consumption.
5. Providing some structure to the 24-hour recall procedures may facilitate data collection. Of course, one must be careful not to lead the answers or bias the response toward anticipated answers.
6. A system of immediate data input to a large data file using computer terminals in the field would facilitate data processing. As the system for collection and analysis is designed, it would be possible to immediately convert field data to food categories and nutrients so the statistical work could proceed rapidly.

References

1. Osteoporosis Consensus Conference. JAMA 252:799–802. 1984.
2. Johnston CC Jr. Studies on prevention of age-related bone loss. In: Peck WA, ed. Bone and mineral research. Amsterdam: Elsevier. 1985.
3. Heaney RP. Prevention of age-related osteoporosis in women. In: Avioli LV, ed. The osteoporotic syndrome. New York: Grune & Stratton. 1983.
4. Melton LJ, Riggs BL. Epidemiology of age-related fractures. In: Avioli LV, ed. The osteoporotic syndrome. New York: Grune & Stratton. 1983.
5. Sowers MR, Wallace RB, Lemke JH. Correlates of mid-radium bone density among postmenopausal women: A community study. Am J Clin Nutr 41:1045–53. 1985.
6. Parfitt AM. Age-related structural changes in trabecular and cortical bone: Cellular mechanisms and biomechanical consequences. Calcif Tissue Int 36:5123–8. 1984.
7. Ruegsegger P, Dambacher MA, Ruegsegger E, et al. Bone loss in premenopausal and postmenopausal women. J Bone Joint Surg 64A:1015–23. 1984.
8. Wahner HW, Dunn WL, Riggs BL. Non-invasive bone mineral measurements. Seminars Nuclear Med 13:282–9. 1983.
9. Mazess RB. Non-invasive methods for quantitating trabecular bone. In: Avioli LV, ed. The osteoporotic syndrome. Detection, prevention, and treatment. New York: Grune & Stratton. 1983.
10. Johnston CC Jr. Non-invasive methods for quantitating appendicular bone mass. In: Avioli LV, ed. The osteoporotic syndrome. Detection, prevention, and treatment. New York: Grune & Stratton. 1983.
11. Darn SM, Rohmann CG, Wagner B. Bone loss as a general phenomenon in man. Fed Proc 26:1729–36. 1967.
12. Chinn HI. Effects of dietary factors on skeletal integrity in adults: Calcium, phosphorus, vitamin D and protein. Bethesda, Maryland: Federation of American Societies for Experimental Biology. 1981.
13. Heaney RP, Gallagher JC, Johnston CC, et al. Calcium nutrition and bone health in the elderly. Am J Clin Nutr 36:986–1013. 1982.
14. Avioli LV. Calcium and osteoporosis. Ann Rev Nutr 4:471–91. 1984.
15. Darn SM, Hawthorne VM. Calcium intake and bone loss in population context. In: Rubin RP, Weiss GB, Putney JW, eds. Calcium in biological systems. New York: Plenum Press. 1985.
16. Raisz LD, Kream BE. Regulation of bone formation. N Engl J Med 309:29–35, 83–9. 1983.
17. Parfitt AM. Dietary risk factors for age-related bone loss and fractures. Lancet 2(8360):1181–4. 1983.
18. Yano K, Heilbrun LK, Wasnich RD, et al. The relationship between diet and bone mineral content of multiple skeletal sites in elderly Japanese-American men and women living in Hawaii. Am J Clin Nutr 42:877–88. 1985.
19. Freudenheim JL, Johnson NE, Smith EL. Relationships between usual nutrient intake and bone-mineral content of women 35–65 years of age: Longitudinal and cross-sectional analysis. Am J Clin Nutr 44:863–76. 1986.
20. Sempos CT, Johnson NE, Smith EL, et al. Effects of intraindividual and interindividual variation in repeated dietary records. Am J Epidemiol 1121:120–9. 1985.
21. Beaton GH, Milner J, McGuire V, et al. Source of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 37:986–95. 1983.
22. Matkovic V, Kostial K, Simonovic I, et al. Bone status and fracture rates in two regions of Yugoslavia. Am J Clin Nutr 32:540–9. 1979.
23. Reed D, McGee D, Yano K, et al. Diet, blood pressure and multicollinearity. Hypertension 7:405–10. 1985.
24. Morgan KJ, Johnson SR, Rizek RL, et al. Collection of food intake data: An evaluation of methods. J Am Diet Assoc 87:888–96. 1987.
25. Johnson NE, Sempos CT, Elmer PJ, et al. Development of a dietary intake monitoring system for nursing homes. J Am Diet Assoc 80:549–57. 1982.
26. Byers T, Marshall J, Fiedler R, et al. Assessing nutrient intake with an abbreviated interview. Am J Epidemiol 122:41–50. 1985.
27. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65. 1985.
28. Chu SY, Kolonel LN, Hankin JH, Lee J. A comparison of frequency and quantitative dietary methods for epidemiologic studies of diet and disease. Am J Epidemiol 119:323–34. 1984.
29. Hankin JH, Kolonel LN, Hinds MW. Dietary history methods for epidemiologic studies: Application in a case-control study of vitamin A and lung cancer. J Natl Cancer Inst 73:1417–22. 1984.
30. Johnson NE, Nitzke S, VandeBerg D. A reporting system for nutrient adequacy. Home Econ Res J 2:210–21. 1974.

Chapter 6 Dietary methodologies for food and nutrition monitoring

by Elizabeth A. Yetley, Ph.D., R.D., Arletta M. Beloian, and Christine J. Lewis, Ph.D., R.D., Center for Food Safety and Applied Nutrition, Food and Drug Administration

Introduction

Since the early 1900's, when the existence of frank nutritional diseases was identified in U.S. population groups, the American public and national policymakers alike have been concerned with the role of nutrition and dietary intakes in public health (1). National monitoring of food consumption began in 1936 with implementation of the continuing series of Nationwide Household Food Consumption Surveys by the U.S. Department of Agriculture (USDA). In the 1960's nutrition monitoring was expanded to include data on individuals with completion of the first Nationwide Food Consumption Survey (NFCS) by USDA and the launching of the Ten State Nutrition Survey by the Department of Health, Education, and Welfare, supported by recommendations from the 1969 White House Conference on Food, Nutrition, and Health. The development of such new programs as the Food Stamp program; food delivery programs for the elderly; the Women, Infants, and Children (WIC) supplemental food program; and nutrition labeling had their genesis from these new concerns and data bases. Nutrition monitoring activities continued through the 1970's, with the completion of the second NFCS by USDA and two National Health and Nutrition Examination Surveys (NHANES) by the National Center for Health Statistics (NCHS). During the 1980's monitoring activities were more focused, with the conduct of Hispanic HANES (HHANES) by NCHS and the Continuing Survey of Food Intakes by Individuals (CSFII) conducted by USDA.

Recently, Federal agencies have paused to evaluate and learn from the flurry of data collection activities of the previous two decades. Critical evaluations of issues involved in the collection and use of dietary assessments were made by several expert committees under the auspices of the National Academy of Sciences (2–4). Currently, an expert panel has been convened under a Food and Drug Administration contract with the Life Sciences Research Office of the Federation of American Societies for Experimental Biology to provide guidelines on the appropriate uses of dietary survey data. Ongoing legislative activities to formalize a comprehensive national nutrition monitoring system also provide evidence of the continuing interest and perceived need by Congress to develop a strategy for improving the nutrition monitoring system in the United States.

When considering nutrition monitoring activities, it is important to recognize that their orientation and the role of dietary data are shifting. Initial concerns with frank nutritional deficiencies have been overshadowed by issues of food safety, questions about the potential for adverse effects of excessive consumption of food components and nutrients, increasing interest in identifying the possible role of dietary factors in prevention of chronic diseases, and new emphasis on developing educational programs relative to health promotion and disease prevention. Although continued monitoring for nutritional inadequacies is needed because of the constantly changing nature of the food supply and concerns about the availability of food for certain populations such as low-income groups, there is a general sense that nutritional deficiencies tend to be less severe than in the past and largely are limited to a few target groups. Furthermore, as the national nutrition monitoring system has expanded its data base to include other types and sources of data (for example, medical and biochemical data, surveys on vitamin-mineral supplement use, screening programs in State health units), the role of dietary intake data is changing. Dietary data are often no longer the primary source of information but are now one of many types of data used for evaluating nutritional and food safety risks and for developing programs and policies.

Characteristics of food and nutrition monitoring

To provide a framework for discussion and recommendations on food and nutrition monitoring, we will borrow concepts presented by Habicht (5–7). We will consider the topic of nutrition monitoring in its broadest sense, which includes not only the traditional assessment of inadequate and normal intakes of food components that protect health but also information on imbalances among food constituents and excessive intakes of food components deleterious to health. This definition expands the use of food intake data from its traditional application for evaluation of the potential risk of nutritional inadequacies to encompass the rapidly increasing need for assessments of dietary exposures to contaminants, toxins, food additives, and nonfood sources of nutrients as well as newer concerns of excessive intakes of food components, as described by the Government pamphlet "Nutrition and Your Health: Dietary Guidelines for Americans" (8). For these reasons, we must consider the topic at hand to be one of food and nutrition monitoring rather than simply nutrition monitoring.

Habicht also notes that the collection of food and nutrition data has a number of public health uses. Each use has its own requirements for data collection and analysis activities. Several of these purposes will be briefly discussed to: (a) clarify how use of dietary data for food and nutrition monitoring compares and contrasts with data needs for epidemiological applications and with other public health data collection activities in nutrition, discussed in previous workshops, and (b) emphasize the need for linkages to national food and nutrition policy and to program activities.

Five basic types of data collection and end-use activities can be described, as follows.

Assessment—The measurement and description of the nutritional status and food intake of a population; the description and prevalence of nutritional or food intake problems; and identification of the subgroups at greatest risk of inadequate or excessive intakes of food components.

Monitoring—The estimation of changes in "true" prevalence of food or nutritional intake problems over time. Monitoring is linked to assessment in that monitoring consists of repeated assessments over a period of time. This type of data use imposes the difficult requirement that repeated assessments be comparable; that is, sampling and measurement methods used in different surveys must be identical or be capable of being adjusted. Then differences between surveys can be attributed with confidence to differences in food or nutrient intakes and not be confounded by artifactual differences resulting from varying methodologies. Ideally, monitoring activities are tied to the development and evaluation of national public health policies and programs; however, they are not designed to promote immediate actions or interventions.

Surveillance—A survey of nutritional or food-related disorders of public health concern for which rapid and effective intervention is needed. This type of data collection is carried out in an ongoing, continuous system intimately linked to an active public health program and is usually focused on high-risk groups and temporally close or well-defined associations. It typically includes a mechanism for rapid intervention and evaluation. Surveillance is often linked to assessment and monitoring through their identification of the food components and nutrition-related disorders appropriate for inclusion in a surveillance system.

Screening—A system that identifies specific individuals for nutritional or public health intervention, often at the community level.

Epidemiological investigation—A comparison of differences in dietary intakes among population groups or subgroups and the relationship of these dietary differences to different risks of health problems. This type of analysis attempts to extrapolate findings of diet-health relationships from one population group to another. Whereas monitoring compares prevalences over a period of time to identify changes in similar population groups, epidemiological research compares differences in dietary intakes between populations with differences in their prevalence or risk of disease. As a consequence, relative differences, rank-ordering, or categorizing techniques for classifying persons as having high or low dietary intakes and high or low risk or rate of disease often are adequate for epidemiological comparisons. In contrast, the prevalence estimates based on "true" or actual intakes are needed in assessment and monitoring activities. However, true prevalences can become a concern for epidemiologists attempting to estimate the public health importance of a diet-health relationship or evaluating how a change in risk factors resulting from intervention has affected the prevalence of a disease. In these cases, the data needs of the epidemiologist are similar to those needed for monitoring purposes.

From here on, we have limited our discussion to the assessment and monitoring types of activities. Surveillance and screening activities, although important, are generally outside the scope of a survey such as NHANES and are more appropriately handled by programs such as those conducted by the Center for Chronic Disease and Health Promotion, Centers for Disease Control, and by local, State, and Territorial public health agencies. Epidemiological applications have been well described by several others participating in this workshop.

In addition to discussing the data needs for assessment and monitoring activities, we will also outline the linkages of these data bases to public health programs. Implicit in food and nutrition monitoring is the concept that the design and evaluation needs of public health programs should be a planned component of these surveys (5). To accomplish this, we need not only information on demographic and descriptive factors to identify subgroups at need but also adequate information on patterns of food use (including information on usual and extreme dietary intakes) and information on how foods are combined within eating occasions, as well as appropriate and sensitive measures for determining a program's progress toward meeting its objectives.

Issues in assessment and monitoring

As outlined above, the primary purposes of assessment and monitoring of data activities are to estimate the prevalence of persons at risk because of inadequate, excessive, or imbalanced intakes of foods or food components and to compare changes in these prevalences over time. The most recent and comprehensive review of issues involved in food and nutrition monitoring can be found in the 1986 report from a National Academy of Sciences (NAS) subcommittee (4), convened for the purpose of evaluating the assessment of nutrient adequacy from food consumption surveys. Using this report and other sources as a basis, we have selected certain key issues for more detailed discussion.

Systematic bias

All measurements have components of random error and systematic bias. Of course, all errors should be minimized, but systematic errors are more serious than random types of error when estimating the prevalence of high or low dietary intakes (4). Systematic error shifts population distribution curves to the left or right. Therefore, estimates of prevalence are directly affected. Furthermore, correction of systematic error from one survey to the next can produce a false and misleading appearance of a real change from one time period to another.

Examples of several potential sources of systematic error with relevance to the NHANES dietary data collection procedures are given below in order to better illustrate this problem. These examples are not comprehensive and are, by necessity, taken out of context relative to the overall survey data base.

Inadequate laboratory methodologies and procedures—A major potential source of systematic error is in food composition data bases and data handling procedures. Significant errors in estimates of intake of food components can occur when laboratory methods for analysis are not good or were not properly performed. An example of the potential bias introduced by this problem can be shown by data from Wolf (9). In this study, the analyzed nutrient content of diets of 22 adults was compared with calculated results using USDA Handbook Number 8. Good agreement was observed for calculated versus collected food weights and estimates of calcium intakes, suggesting that the collection and recording procedures were consistent and that reference table values for calcium were accurate. However, results for other nutrients were quite variable. Iron estimates were biased toward high values for the method by calculation compared with the method by analysis; zinc estimates were biased toward lower values for the method by calculation. Using the iron content of beef and pork from several data bases (10–12) as an additional example, the impact on standard reference values when food composition data are corrected is shown in table 1. Different iron values were used for the same food item in several surveys, depending upon the date of the available food composition data base. Thus, the prevalence of persons with low intakes of iron will appear to change over time (assuming that no other changes in intake patterns or composition values have occurred). In evaluating monitoring data, it is important, therefore, that changes resulting from corrections in the original reference composition data base not be misconstrued as resulting from real changes in nutrient intakes in subsequent surveys.

Food descriptors—The completeness and adequacy of food descriptions, the appropriate coding of data, and the changes in descriptors used over a period of time constitute a second major potential source of systematic bias. Again, several examples can be given.

Table 1. Iron content of beef and pork from data bases used during National and Hispanic Health and Nutrition Examination Surveys

Food and survey	*Food code*	*Milligrams of iron per 100 grams of food*
Beef, ground, cooked, lean		
NHANES I, 1971–74[1]	00368	3.50
NHANES II, 1976–80[2]	00368	3.50
HHANES, 1982–84[3]	215–0120	2.18
Pork chops, cooked, lean and fat		
NHANES I, 1971–74[1]	01717	3.40
NHANES II, 1976–80[2]	01717	3.40
HHANES, 1982–84[3]	221–0101	0.93

[1]First National Health and Nutrition Examination Survey. SOURCE: National Center for Health Statistics. National Health and Nutrition Examination Survey 1971–75 public use data tape documentation: Model gram and nutrient composition. Tape nos 4702 and 4703. Hyattsville, Maryland. 1981.
[2]Second National Health and Nutrition Examination Survey. SOURCE: National Center for Health Statistics. National Health and Nutrition Examination Survey1976–1980 public use data tape documentation: Model gram and nutrient composition. Tape nos 5702 and 5703. Hyattsville, Maryland. 1982.
[3]Hispanic Health and Nutrition Examination Survey. SOURCE: Human Nutrition Information Service. Manual of food codes for individual intake. Hyattsville, Maryland: U.S. Department of Agriculture. 1985.

First, specificity and consistency in the descriptors of foods are important, particularly for those processed foods in which nutrient and ingredient compositions vary significantly by brand, fortification or food processing practices may result in significant differences among brands or within a given brand over time, and brand loyalty by consumers is likely to occur. For instance, several sweetened cola drinks are listed under a single food code in the manual of food codes now used jointly by USDA and NCHS (12). Although these beverages have similar nutritional profiles, their profiles for other food components may vary considerably. As shown in table 2, caffeine content (13) or a theoretical food additive ("Compound X") could represent such food components. Loss of the

Table 2. Food component content of cola soft drinks

Soft drink, cola type (924–1031)[1]	*Components per 369-gram serving*		
	Kilocalories [2]	*Caffeine (milligrams)*[3]	*Compound "X" (micrograms)*
Coca-Cola	144.0	45.6	...
Canned	...	...	92
Not canned	...	...	–
Pepsi Cola	157.2	38.4	...
Canned	...	...	–
Not canned	...	...	–
RC Cola	155.6	36.0	...
Canned	...	...	–
Not canned	...	...	–
Jamaica Cola	165.0	30.0	...
Canned	...	...	129
Not canned	...	...	–
Average value	155.4	37.5	28
"Maximum" value	165.0	45.6	129

[1]Code numbers 924–1031. SOURCE: Human Nutrition Information Service. Manual of food codes for individual intake. Hyattsville, Maryland: U.S. Department of Agriculture. 1985.
[2]National Soft Drink Association. Washington, D.C.: personal communication. 1981.
[3]Institute of Food Technologists' Expert Panel on Food Safety and Nutrition. Caffeine. Food Technol 37:87–91. 1983.

specific brand and packaging identities for the individual food items in this coding category will require that assumptions be made in using these data for estimating intakes of caffeine or the theoretical Compound X. Averages for the components can be used; however, averages will obviously underestimate intakes of "heavy" consumers who not only consume the product several times a day but may also have a brand loyalty to the product with the highest concentrations. Furthermore, bias will be introduced if some population subgroups use brand products in differing proportions from other groups. An occasionally used "conservative" alternative for food safety evaluations is to assume that all foods within the food category contain the component of interest and that its concentration is at maximum level. This approach would clearly overestimate intakes (table 2). Yet, although the incorporation of more specific food descriptors into assessment and monitoring activities is desirable, it would also cause a "correction" in the data base. Therefore, comparison of intakes based on detailed brand information in subsequent surveys with the grouped data used in the past would confound the ability to differentiate between true changes in intake and changes because of the artifact of improving the reference data base.

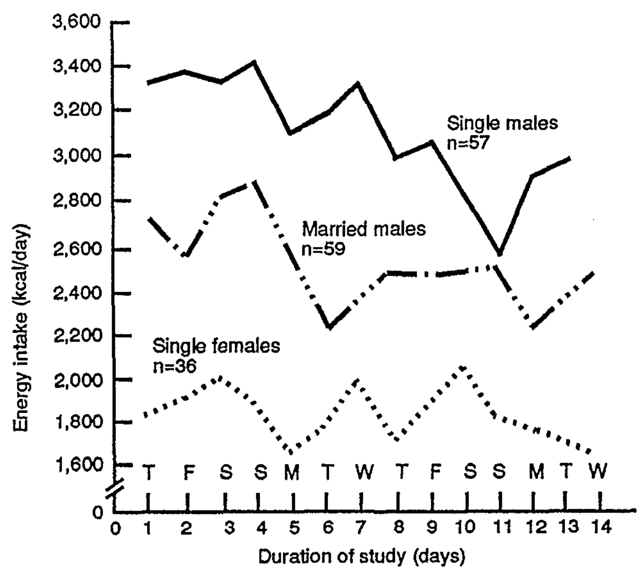
A second example of potential sources of systematic error caused by changes in data handling can be seen by reviewing the coding structure for green beans from the first NHANES survey to the most recent one, based on the current coding system used jointly by NFCS-CFSII and HHANES. Changes in assumptions about the processing of green beans were apparently made from NHANES I to NHANES II, as evidenced by changes in sodium content of codes 183 and 184 (table 3). (In general, canned

Table 3. Descriptions and composition of green beans used during Health and Nutrition Examination Surveys

Survey, food code, and description	*Fat*	*Sodium*
NHANES I, 1971–74[1]	Grams per 100 grams of food	Milligrams per 100 grams of food
182: Beans, snap, green, raw	0.2	7
183: Beans, snap, cooked in small amount of water, short time	0.2	4
184: Beans, green, cooked (drained)	0.2	4
185: Beans, snap, canned, regular pack, solids and liquids	0.1	236
186: Beans, green, canned, solids only	0.2	236
187: Beans, snap, canned, regular pack, drained liquid	0.1	236
188: Beans, snap, special dietary pack, solids and liquids	0.1	2
189: Beans, snap, special dietary pack, drained solids	0.1	2
190: Beans, snap, special dietary pack, drained liquid	0.1	2
191: Beans, snap, frozen, cut, not thawed	0.1	1
192: Beans, snap, frozen, cooked, boiled, drained	0.1	1
193: Beans, French style, not thawed	0.1	2
194: Beans, snap, French style, cooked, boiled, drained	0.1	2
204: Beans, snap, frozen, cooked, boiled, drained	0.1	1
NHANES II, 1976–80[2]		
182: Beans, snap, green, raw	–	7
183: Beans, snap, cooked in small amount of water, short time	–	240
184: Beans, green, cooked in large amount of water, long time	–	240
185: Beans, snap, canned, regular pack, solids and liquids	–	236
186: Beans, green, canned, solids only	–	236
187: Beans, snap, canned, regular pack, drained liquid	–	236
192: Beans, snap, frozen, boiled, drained	–	236
194: Beans, snap, French style, boiled, drained	–	238
204: Beans, snap, frozen, boiled, drained	–	237
HHANES, 1982–84[3]		
752-0500: Beans, string, NS as to color or added fat	3.1	253
752-0501: Beans, string, green, cooked, NS as to fat added in cooking	3.1	253
752-0502: Beans, string, green, cooked, fat not added in cooking	0.3	234
752-0503: Beans, string, green, cooked, fat added in cooking	3.1	253
752-0511: Beans, string, green, canned, low sodium, cooked, NS as to fat added in cooking	3.0	35
752-0512: Beans, string, green, canned, low sodium, cooked, fat not added in cooking	0.1	2
752-0513: Beans, string, green, canned, low sodium, cooked, fat added in cooking	3.0	35

[1]First National Health and Nutrition Examination Survey. SOURCE: National Center for Health Statistics. National Health and Nutrition Examination Survey 1971–75 public use data tape documentation: Model gram and nutrient composition. Tape nos 4702 and 4703. Hyattsville, Maryland. 1981.
[2]Second National Health and Nutrition Examination Survey. SOURCE: National Center for Health Statistics. National Health and Nutrition Examination Survey 1976–80 public use data tape documentation: Model gram and nutrient composition. Tape nos 5702 and 5703. Hyattsville, Maryland. 1982.
[3]Hispanic Health and Nutrition Examination Survey. SOURCE: Human Nutrition Information Service. Manual of food codes for individual intake. Hyattsville, Maryland: U.S. Department of Agriculture. 1985.

NOTE: NS is not specified.



SOURCE: Schnakenberg et al. (14, p. 188). Reprinted with permission from *Assessing Changing Food Consumption Patterns*, 1981, National Academy Press, Washington, D.C.

Figure 1. Energy intakes by sex and marital status

vegetables tend to be high in sodium, as salt is added during processing; fresh or frozen forms tend to be naturally low in sodium.) Codes 183 and 184 in NHANES I and II illustrate the difficulty in estimating both the intakes of an added nutrient like sodium and the changes in sodium intake over a period of time, because it was apparently assumed in NHANES II that sodium is added by all respondents prior to eating the product (unless a low-sodium canned product is specified). Thus, a respondent who uses fresh or frozen green beans without added salt will appear to have consumed the salted product. In contrast, for NHANES I, no added salt was assumed.

Additional changes are seen as we move from NHANES II to HHANES, which used the joint coding system with USDA. Here, not only are food descriptions different from those used in previous HANES, but also fat added during cooking is included in the green bean code rather than listed separately. If such a change in code is followed for other vegetables, there will be an apparent change in sources of fat between NHANES II and future surveys that use the new coding system.

Specific population groups—Another concern in attempting to identify changes in prevalence of risk over time is the possibility of systematic bias in reporting food intake for an entire subgroup. There are insufficient data to evaluate this issue directly, but several pieces of evidence suggest that this may be a problem. For example, suggestions of possible sex and age differences in reliability of reporting of dietary results have been noted (4). Schnakenberg et al. (14) found quite different consistencies in reporting patterns between married and single persons for both males and females over a 14-day collection of dietary records. Specifically, as shown in figure 1, respondent fatigue is suggested by decreasing levels of reported energy intake for single men; married men and

Table 4. Estimated intakes of calcium and iron by infants from 3 surveys

Variable	*SMIF, 1974–80*[1]	*NHANES II, 1976–80*[2]	*NFCS, 1977–78*[3]
Age (months)	6	6–11	0–12
Kilocalories	880	880	880
Calcium (milligrams)	827–1,009	823	877
Iron (milligrams)	4.7–10.2	11.4	19.3

[1]SMIF = Selected Minerals in Foods.
[2]NHANES = National Health and Nutrition Examination Survey.
[3]NFCS = Nationwide Food Consumption Survey.

SOURCE: Pennington JAT, Wilson DB, Newell RF, et al. Selected Minerals in Foods Surveys, 1974 to 1981/82. J Am Diet Assoc 84:771–80. 1984.

single women showed less attenuation over time. In another study, plate waste was compared with recall data in hospitalized patients (15). Results showed that elderly subjects were more likely than young adults to underreport their caloric intake. Probing by interviewers resulted in increased estimates of reported intakes. Hallfrisch, Steele, and Cohen (16) found that adult males and females both tended to underreport their caloric intakes, but the deficit was more severe in females than in males (35 percent and 15 percent, respectively). In short, this evidence suggests that there may be systematic reporting errors that vary from group to group, but the data are limited and difficult to interpret.

Bias in estimating dietary intakes of specific subgroups is more likely to occur when the population group of interest consumes a limited variety of foods and when highly processed foods comprise a large portion of the diet. An example of this potential problem can be illustrated by comparing estimates of dietary intakes of infants obtained from three surveys (table 4). The Selected Minerals in Foods (SMIF) data provide, for foods purchased in 1974–80, nutrient values based on a diet representing infant eating patterns in the 1965 NFCS (17). After adjusting the NHANES II (1976–80) and the 1977–78 NFCS results to the kilocalorie level of the SMIF diets, it can be seen that estimates of calcium intake among all three data bases were similar, but iron intakes varied considerably. Because a limited number of foods are available for infants (that is, infant formulas and infant cereals constitute the major source of nutrition in infant diets), intakes are particularly sensitive to variations in fortification practices for the foods used and also to the particular foods selected. Estimates of total intakes can thus vary widely. Because the three data bases in table 4 differed in their original sample population, the specificity of food descriptions used, the assumptions as to whether fortified or nonfortified products were chosen, and composition values for the iron content for supposedly similar infant foods, it is not surprising that quite different intake estimates were obtained.

Underreporting of nutrients, foods, or food groups—Finally, if accurate estimates of intakes are to be made, it is extremely important that all sources of consumed food and nutrients be identified, quantified, and incorporated into the intake estimates. Alcoholic beverages, minerals and electrolytes supplied by water and medications, organ meats, vitamin-mineral and other food supplements, and seasonal foods such as melons should be included. The low estimates of consumption of alcoholic beverages compared with sales data have been noted (4) and raise concerns not only about the accuracy of intake estimates for this beverage group but also about the potential for underestimating energy intakes. Water can be a significant source not only of minerals and electrolytes but also of contaminants. Adequate intake estimates and home use of water for reconstituting beverages are not available from any source at the present time. The potential for medications to be a significant source of nutrient minerals is suggested by recent advertisements for using a popular antacid as a calcium source. Use of organ meats, although important both nutritionally and toxicologically, is hard to accurately estimate with a limited number of days of dietary data because of infrequent use. An advantage of obtaining information on vitamin-mineral supplement use, besides the obvious one of more complete intake estimates, is suggested by empirical data from Sempos et al. (18), who found that when dietary intake estimates included intakes from both supplements and foods, the ratio of intraindividual to interindividual variation in variance estimates decreased significantly and in the desired direction for reducing the measurement error in estimating "usual" intakes.

Usual intake

Another problem in attempting to identify population groups at greatest risk is that in order to interpret dietary assessments, one needs to be able to estimate usual intakes as accurately as possible. This issue has been reviewed by others (3,4,18) and discussed by previous speakers in this workshop. Estimations of usual intakes must be derived from multiple days of data in order to reduce the effect on population distributions of normal day-to-day variabilities in intakes by individuals. An example of how failure to account for intraindividual variation in food safety evaluations may affect findings can be illustrated by the data in tables 5 and 6. These results (19) for children ages 2–5 years show that quite different intake estimates for the "heavy" consumer (90th percentile) can be obtained, depending on whether results are expressed as single-day estimates or as 14-day averages for each respondent (table 5). Furthermore, estimates of the percentage of children ages 2–5 years exposed to the food component of interest may vary, depending on the number of days of information per subject (table 6).

As noted by Hegsted (20) and as discussed by the NAS subcommittee (4), when determining usual intakes in populations, the need is not for reliable estimates of dietary intake for each person but rather for reliable estimates of the distribution of usual intakes for the population.

The NAS subcommittee indicated that, unlike individual intakes, the distribution of usual dietary intakes for

Table 5. Single-day and 14-day average estimates of monosodium glutamate consumption among children ages 2–5 years who consume monosodium glutamate

Estimate	*Mean*	*90th percentile*
	Milligrams per day of monosodium glutamate	
Single day	99	379
14-day average	87	176

Extracted from ref. (19), p.187, by courtesy of Marcel Dekker Inc.

Table 6. Single-day and 14-day average estimates of the percent of children ages 2–5 years consuming chondrus extract

Estimate	*Percent consuming chondrus extract*
Single day	53
14-day average	95

Extracted from ref. (19), p.188, by courtesy of Marcel Dekker Inc.

the population can be approximated from a modest number of repetitions of the 1-day intake data. However, they pointed out that seasonal variations in intake and variation between weekdays and weekend days must also be taken into account in the data collection and that the replicated observations should be independent of one another in time rather than taken on consecutive days. The subcommittee further suggested that 3 days of observation may be more than is required for the derivation of distribution of usual intakes. Whether this recommendation holds for nonnutritional food components that are not ubiquitously distributed or consumed at least in some minimal amount on a daily basis is not clear.

Links to public health programs

Clearly, a major role for food and nutrition monitoring is its linkage to the development and evaluation of national public health policies and programs (figure 2). At a minimum, these data should be capable of describing the adequacy and safety of the U.S. diet and of identifying problem intakes and population groups at greatest risk. In addition, the monitoring data bases should provide enough information to aid in developing strategies for resolving public health problems, for predicting program effectiveness, and for evaluating program impacts. All of these uses require, in addition to information on total intakes of food components, data on eating habits and types of food consumed. Supportive evidence is also needed to assure that the definition of the problem and intervention are appropriate, for instance, that a perceived public health problem is in fact related to food habits and not caused by other factors. Evidence is also needed to show that recommended changes in food habits will not carry the risk of unintended adverse health outcomes.

Certain specific program needs (both design and evaluation needs) requiring food and nutrition monitoring have been described elsewhere (3) and are shown in table 7 in order to illustrate the diversity of need for dietary data. In addition, programs such as food assistance and nutrition education also have important applications

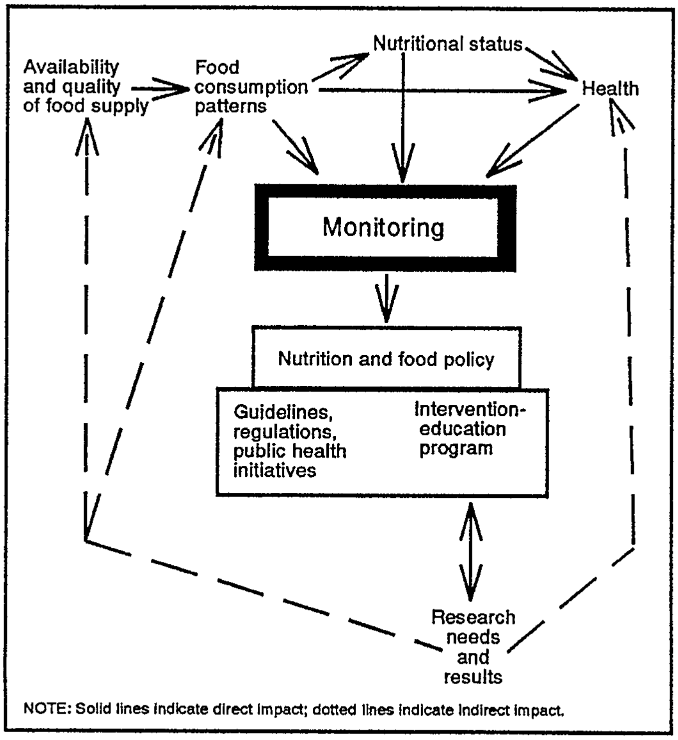


Figure 2. Linkages: Monitoring and national policy

Table 7. Selected program needs requiring food and nutrition monitoring

Program or activity	*Agency involved*
Nutrient fortification of the U.S. food supply	FDA
Health education and information programs:	
Cholesterol Education Program	NIH
Blood Pressure Education Program	NIH
Nutrition labeling initiatives	FDA
The 1990 Nutrition Objectives for the Nation	DHHS
Iron status in pregnancy	
Growth and development in children	
Overweight and obesity in adults	
Serum cholesterol	
Daily sodium intakes	
Review of the Recommended Dietary Allowances	NAS
Safety of the food supply	FDA, EPA, USDA
Food additives, contaminants, toxins, other nonnutrient components (includes Total Diet Study)	
Acute adverse reactions to foods	
Interactions among food components	
Excessive intakes of nutrients	
Monitoring of prevalence of specific diet-related diseases and related dietary intakes	NIH, FDA, DHHS
Osteoporosis	
Cancer	
Cardiovascular disease	
Autoimmune deficiency disease syndrome	
Alzheimer's disease	
Hypertension	

NOTE: DHHS is U.S. Department of Health and Human Services. EPA is Environmental Protection Agency. FDA is Food and Drug Administration. NAS is National Academy of Sciences. NIH is National Institutes of Health. USDA is U.S. Department of Agriculture.

for dietary data but are not listed in the table because traditionally their design and evaluation have been more closely linked to NFCS than to NHANES.

The linkage of public health initiatives to the food and nutrition monitoring system requires, in many cases, an additional burden for data collection over and above the capability of assessing intakes of food components. Specifically, it requires identification of food and meal patterns, of how much and which types of food are consumed together in single eating occasions, and of changes in eating patterns that reflect both quantitative and qualitative modifications. An example of the last case is the ability to detect a change in usage of types of foods that differ only in processing, for example, switching from canned to frozen vegetables with a concomitant change in sodium or other food additive intake. Identification of the types of foods consumed together by eating occasion has relevance to issues of bioavailability of food components (nutrients, toxins, and so forth) and other possible interactions, such as that of aspartame with high-carbohydrate, low-protein meals or snacks.

Data on eating habits and patterns are as important as assessment of total dietary intakes. The uses of such data require detailed and complete information on types of foods consumed, on foods combined within an eating occasion, and on sources of food components. Similar statistical and data collection issues arise as with the estimation of "usual" dietary intakes of food components. As noted by the recent NAS subcommittee (4), the impact of individual day-to-day variations is a problem in descriptive data on patterns of food use, just as it is for estimates of total dietary intakes of food components. In fact, estimations of food patterns may require more replications (number of days of dietary data) per individual than is needed for estimating "usual" distributions of intakes for population groups.

Recommendations

To this point, we have discussed the characteristics of a food and nutrition monitoring program and have identified data needs and potentially serious sources of systematic bias. By definition, the monitoring system measures changes over time. Such measurement requires baseline data and comparable assessments with subsequent surveys. As summarized in table 8, both NHANES and NFCS have used 24-hour recall methodologies to collect dietary data. These two data systems, however, have differed both within one system from survey to survey and between the two systems. Such differences could affect results, although there is no way to adequately or easily assess the magnitude or direction of such effects if they have occurred. Our recommendations, therefore, are oriented not only to improving the types and program utility of information from NHANES, but are also influenced by concerns that the data collection process be more thorough so as to better monitor time trends in the eating habits of the U.S. population and to compare results across surveys.

Methodology: 24-hour recall

We recommend that the 24-hour recall methodology be continued. This methodology provides continuity with

Table 8. Major characteristics of food consumption surveys: National Health and Nutrition Examination Survey (NHANES) and Nationwide Food Consumption Survey (NFCS)

Characteristic	NHANES I, 1971–75	NHANES II, 1976–80	NFCS, 1965	NFCS, 1977–78
Sampling frame	Individuals	Individuals	Household	Household
Sample size	30,000	20,000	15,000	30,700
Season	Follows sun	Follows sun	Spring	All days
Days of week	Sunday–Thursday	Sunday–Thursday	All days	All days
Dietary methodology	24-hour recall plus food frequency	24-hour recall plus food frequency	24-hour recall	24-hour recall plus 2-day record
Number of days of data collected	1	1	1	3
Collection method	Interview	Interview	Interview	Interview plus self-administered record
Interview environment	Mobile examination center	Mobil examination center	Home	Home
Serving size estimates	Food models	Food models	Household measures	Household measures
Number of foods in data base	3,400	2,600	1,900	3,700

past NHANES as well as with NFCS and other surveys, within the caveats discussed above. It also has the potential for providing the types of information on eating behaviors needed for program design and evaluation.

Determination of "usual" intake

Multiple and independent replications of dietary intakes are needed to adjust distributions of population groups to represent "usual" intakes. Although the NAS subcommittee (4) recommended 2 days of dietary intake data as being adequate for adjusting total dietary intake values, it also noted that more days may be needed for examining eating patterns. Certainly, 3 days of intake data would be preferable to 2, but questions of cost efficiency and logistics must also be considered.

Replications of days of dietary intake collection would optimally be done by using the same technique on all respondents—for example, bringing all respondents back to the mobile examination center for additional dietary interviews. However, such an approach may pose too great a burden on the survey. It is preferable that the second and subsequent days of dietary intake collection not be conducted in the home of the respondent because, unlike the mobile examination center, the home may have an "audience" of family members who could exert social pressures and possibly influence the information provided by the respondent.

Several alternatives can be proposed. Multiple days of intake on a subsample of the total sample could be obtained. This subsample should be representative of the sex, age, and socioeconomic groups included in the survey and could return to the mobile examination center for all dietary interviews. Results from this subsample could be used to estimate the intraindividual variation in dietary intakes and to adjust the population distributions accordingly. Another possibility is the use of a telephone interview to obtain subsequent 24-hour recalls on subjects. This procedure has the advantage of cost and time effectiveness but does not meet the recommendations for a common data collection method. If this approach is used, respondents need to receive some training at the mobile examination center on how to use paper or household food models for estimating serving sizes during the telephone interview. A third alternative approach would be to use calculations of intraindividual variation from NFCS to adjust the distribution of intakes. However, differences in sampling and data collection between NHANES and NFCS raise serious questions as to the appropriateness of extrapolation based on such corrections.

In addition to providing corrections for intraindividual variations in food intakes, the recall procedures should be based on "independent" rather than consecutive days of intake, be representative of the entire week including weekends, and include seasonal variations as much as possible.

Detailed food descriptors

Detailed and accurate descriptions of food items in the survey data base will improve the usefulness of the data bases. Additionally, evaluation of comparability of data for making time trend estimates will be facilitated.

Because of the complexity of the food supply and eating habits, we recommend that an automated interactive computer system be used by the interviewer. Such a system not only would provide immediate coding of information and timeliness of data release but also could be designed to provide the prompts needed to thoroughly probe for all relevant and obtainable information and to standardize interviewing techniques. The importance of such a system's ability to standardize the interview should not be overlooked. As suggested by Campbell and Dodds (15), differences in probing techniques and followup questions used by interviewers can greatly influence the information provided by the respondent.

NFCS and HHANES are currently using a common reference food composition data base and food coding system (12). This approach has been strongly recommended by past NAS subcommittees (3,4) to provide linkages between the two data systems. We support continued use of a common reference system but recommend that additional data fields be added for each food item to allow, where appropriate and feasible, information on processing and packaging, on brand names, on sources of foods, and on use of home water supplied in food preparation. (This assumes, of course, a computerized automated interview system.) Judgment and reasonableness are needed to determine these extra data fields. For example, most persons can accurately identify the brand of soft drink they consume and indicate whether it comes from a can, a plastic bottle, or a soda fountain, but they are not likely to be able to provide similar information for meats. Vegetables consumed may not be identifiable by brand name, but consumers can usually distinguish between canned and uncanned products. In another example, brands of margarine and cooking fats tend to vary greatly in fatty acid composition. Information on brands of fats and margarine for home use of products could be obtained either in the initial home interview or in followup interviews. For estimating food additive intakes, it is important to know whether the apple pie eaten was prepared from "scratch" at home, purchased frozen in the supermarket and then cooked at home, or eaten at a restaurant. Flavor and color additives vary depending on whether red- or green-colored gelatin desserts are chosen. Such information can be easily obtained during the interview and would greatly expand the usability of the dietary intake data.

Complete intake data

As noted previously, some types of foods and beverages tend to be underreported or are difficult to estimate accurately because of either infrequent use or the need for special probes. These items include, but are not necessarily limited to, alcoholic beverages, water, organ meats, vitamin-mineral and food supplements, sugar and salt substitutes, seasonal foods, and medications that can make mineral contributions to the diet. It is important that sufficient probing be conducted so these items are reflected in the 24-hour recall. In addition, a special frequency questionnaire is needed to specifically estimate intakes over a recent and specified time period. As consumption of some of these items is known to vary depending upon day of the week (higher alcohol consumption on weekends, for instance), such considerations need to be reflected in the frequency questionnaire in order to provide a complete picture of intake. Additionally, because water can be a significant source of minerals and electrolytes as well as certain contaminants, it would be useful to obtain some estimate of frequency and quantity of water consumption as well as to identify foods that are reconstituted with water or have home supplies of water added to them. To provide composition data on drinking water, an aliquot of tap water could be obtained from each of the homes during the initial interview and analyzed for mineral, electrolyte, and contaminant content.

Reliability

The ability to extrapolate results of reliability and validity estimates from small, well-controlled studies to a field survey in which all age and socioeconomic classes are represented is questionable. Also, in comparing two different methodologies for the same individuals, there is the possibility of similarities between methods because of multiple systematic biases that are directionally the same and occur at the same time. Therefore, as much as possible, internal checks should be incorporated into the survey to evaluate the reasonableness of relative intake differences among subgroups. Several approaches could be investigated, including use of such statistical techniques as split half reliability measures to estimate reproducibility of types of information for subgroups and reproducibility of specific foods or food groups. Biological markers are theoretically ideal to check validity of intake data, but in reality such measures are probably not feasible under field survey conditions because they frequently involve multiple and complete 24-hour urine or stool samples. However, similar patterns of differences among subgroups between intake estimates and biochemical indicators that reflect recent intakes (such as serum ascorbic acid) would provide some assurance that dietary intakes are reliable. However, this approach must be used cautiously, as comparisons could be confounded by genetic (racial) and life cycle differences in metabolism of nutrients and by differences among socioeconomic or age groups in infection rates. Also, for many of the serum or plasma levels of nutrients, there is a threshold effect; thus their use is limited largely to checks for low intake rather than high intake. Furthermore, these approaches can provide only qualitative information; they cannot be used to validate the accuracy of quantitative estimates of intakes.

To some extent, the reasonableness of estimates can also be checked externally using other data bases and surveys. For instance, the per capita disappearance data for various food commodities compiled by USDA, as well as food marketing data gathered by the business sector and private organizations, can be compared with the intake values obtained in NHANES. However, the data collection methodologies and final form of the data can differ a great deal from those of NHANES; therefore, the "truth" of the comparison, again, is limited.

Documentation

As mentioned earlier, incompleteness in food description, food intake records, and food composition information can lead to systematic bias. To aid in identifying these types of biases, adequate documentation for dietary data is needed. This documentation should be appropriate for evaluating the completeness and accuracy of the reference food data base. We recommend the following documentation strategies:

- Imputed intakes should be flagged for both food code and serving size information.
- Information provided by a surrogate or proxy should be flagged.
- The environmental condition for subsequent 24-hour recalls should be identified to indicate such pertinent factors as who other than the respondent was present or listening during the interview.
- Imputed values in the food composition reference tables should be flagged so that there is an indication of the completeness of the composition data base and the impact of imputed values on data analysis can be considered.

- Within reason, documentation as to the adequacy and accuracy of the analytical laboratory methods used to determine food composition values should be available to users and should specify (where appropriate) the precision, bias, sensitivity, specificity, levels of contamination or interference, and degree of confidence relative to the analytical method.
- Because NHANES III data will undoubtedly be used for comparisons with past surveys in order to evaluate changes over time, the survey documentation should, at a minimum, incorporate a warning to users as to the need to check compatibility, completeness, and the nature of assumptions for the reference food composition data bases and food descriptors from survey to survey.

References

1. Ostenso GL. National nutrition monitoring system: A historical perspective. J Am Diet Assoc 84(10):1181–5. 1984.
2. National Academy of Sciences, National Research Council, Committee on Food Consumption Patterns. Assessing changing food consumption patterns. Washington: National Academy Press. 1981.
3. National Academy of Sciences, National Research Council, Coordinating Committee on Evaluation of Food Consumption Surveys. National survey data on food consumption: Uses and recommendations. Washington: National Academy Press. 1984.
4. National Academy of Sciences, National Research Council, Subcommittee on Criteria for Dietary Evaluation, Coordinating Committee on Evaluation of Food Consumption Surveys. Nutrient adequacy: Assessment using food consumption surveys. Washington: National Academy Press. 1986.
5. Habicht JP, Lane JM, McDowell AJ. National nutrition surveillance. Fed Proc 37(5):1181–7. 1978.
6. Habicht JP. Some characteristics of indicators of nutritional status for use in screening and surveillance. Am J Clin Nutr 33:531–5. 1980.
7. Habicht JP, Meyers LD, Brownie C. Indicators for identifying and counting the improperly nourished. Am J Clin Nutr 35:1241–54. 1982.
8. U.S. Department of Agriculture and U.S. Department of Health and Human Services. Nutrition and your health: Dietary guidelines for Americans. Home and Garden Bulletin no 232, 2d ed. Washington. 1986.
9. Wolf WR. Assessment of inorganic nutrient intake from self-selected diets. In: Beecher GR, ed. Beltsville Symposium in Agricultural Research: Human Nutrition Research. London: Attanheld, Osmun Co. 1981.
10. National Center for Health Statistics. National Health and Nutrition Examination Survey 1971–75 public use data tape documentation: Model gram and nutrient composition. Tape nos 4702 and 4703. Hyattsville, Maryland. 1981.
11. National Center for Health Statistics. National Health and Nutrition Examination Survey 1976–80 public use data tape documentation: Model gram and nutrient composition. Tape nos 5702 and 5703. Hyattsville, Maryland. 1982.
12. Human Nutrition Information Service. Manual of food codes for individual intake. Hyattsville, Maryland: U.S. Department of Agriculture. 1985.
13. Institute of Food Technologists' Expert Panel on Food Safety and Nutrition. Caffeine. Food Technol 37:87–91. 1983.
14. Schnakenberg DD, Hill TM, Kretsch MJ, et al. Diary-interview technique to assess food consumption patterns of individual military personnel. In: Assessing changing food consumption patterns. Washington: National Academy Press. 1981.
15. Campbell VA, Dodds ML. Collecting dietary information from groups of older people. J Am Diet Assoc 51:29–33. 1967.
16. Hallfrisch J, Steele P, Cohen L. Comparison of seven-day diet record with measured food intake of twenty-four subjects. Nutr Res 2:263–73. 1982.
17. Pennington JAT, Wilson DB, Newell RF, et al. Selected minerals in foods surveys, 1974 to 1981/82. J Am Diet Assoc 84:771–80. 1984.
18. Sempos CT, Johnson NG, Smith GL, et al. Effects of intraindividual and interindividual variation in repeated dietary records. Am J Epidemiol 121:120–30. 1985.
19. Yetley EA, Hanson EA. Data sources and methods for estimating consumption of food components. J Toxicol Clin Toxicol 21(1 and 2):181–200. Marcel Dekker Inc., N.Y. 1983–84.
20. Hegsted DM. Problems in the use and interpretation of the Recommended Dietary Allowances. Ecol Food Nutr 1:255–65. 1972.

First discussant's remarks about monitoring

by Eleanor M. Pao, Ph.D., Human Nutrition Information Service, U.S. Department of Agriculture

The paper by Yetley, Beloian, and Lewis is excellent. The point that "food" as well as "nutrition" monitoring should be a topic under consideration is very appropriate. As stated, alleviation of deficiency diseases and introduction of food programs, such as the Food Stamp program, introduced in 1961, did not eliminate hunger nor ensure that all persons consumed adequate diets. Thus, even in an affluent country such as the United States, the ability of residents to be well nourished cannot be taken for granted. Consequently, in recent years, the nutrition monitoring approach has gained support as a valuable adjunct to the conduct of large-scale, periodic assessment surveys. The latter are too cumbersome to discover nutritional problems which may turn up suddenly in particular population groups. However, the development of the envisioned National Nutrition Monitoring System (NNMS) to include multiple types and sources of data in addition to dietary intake data presents some real challenges.

Linkage among surveys and other data bases in NNMS

First, it would seem that a much higher priority must be assigned to development of linkages among the various data sets or surveys in the monitoring system. In fact, I would emphasize and extend the mention of necessity for linkage between the data bases (including surveys) and the public health programs by urging linkage among the surveys and the data bases as well. A system of linkage among surveys of agricultural production, food consumption, nutritional status, health, and disease prevalence and other data bases should begin at the planning and design stages in order to obtain all the necessary information for linking.

Comparability between NHANES and NFCS

A National Health and Nutrition Examination Survey (NHANES) and Nationwide Food Consumption Survey (NFCS) working group recently studied a number of variables common to both surveys and found a few that were identical. However, modification of a few more questions and definitions and addition of some new questions would contribute further to linkage between the two surveys and thus provide for increased use of both sets of data. These include mostly socioeconomic, demographic, household, and personal characteristics. Some variables seem to be locked into dissimilar definitions. For example, NHANES and NFCS designate a reference person and a head of household (male head, if present; otherwise female head), respectively, around whom relationships in the household are specified. In practice, the reference person and the head of household often are the same person. But if the household organization does not permit ready identification of this person, the two surveys differ in defining to whom that role should be assigned.

One of the most important underlying reasons for differences is the uniqueness of purpose of each survey. This is manifested in the sample designs. For example, the samples of households in NHANES and NFCS differ on the definitions of eligible group-quarter households, with NHANES including some which NFCS excludes.

A common core of questions on food programs would appear to be a reasonable goal for the two surveys in 1987 and 1988. It would be useful in both surveys to have an identical question to determine the respondent for children under 12 years and other persons not answering for themselves. Surrogate reporting needs better documentation and examination. Definitions and protocols for acquisition of information on food intakes, dietary practices, and other dietary factors have not yet been studied for possible linkage.

The possibility of one or two questions about health that could be used to further link NHANES and NFCS might be explored in order to expand the usefulness of the two data bases in the nutrition monitoring program. Both surveys now include a general question on the respondent's assessment of his or her own health—excellent, very good, good, fair, or poor. However, with the public's interest in and media coverage of the role of dietary intake in the development of chronic diseases, it would seem likely that respondents could self-report whether or not they have certain conditions, such as heart conditions, hypertension, cancer, diabetes mellitus, osteoporosis, and dental problems. The respondent-reported prevalence could be compared with the physician-determined prevalence, which could then be compared with self-reports in subsequent surveys.

Age groups to be surveyed

The age groups to be included in NHANES III may not have been determined yet, but there seems to be a significant lack of data on the elderly beyond the age of 74 years. The participation of these individuals in some phases of NHANES III might be limited, but there seem to be many reasons for including this older generation in the dietary intake part of the survey. Foremost, this age group is a steadily increasing percentage of the population. Health care costs of the elderly appear to be higher than those for any other age group. Diet is important in maintaining vigor and health. Dietary management is often part of health management and containment of health costs. The elderly are vulnerable to chronic diseases that are related to dietary factors. They may have problems with obesity. They may also be on reduced incomes, which means that diets must be planned more carefully to maintain good nutrition. People in this age group may be vulnerable to food fads or fad diets in a search for better health as well as being heavy users of vitamin, mineral, and other dietary supplements. For these and other reasons, the relationships of diet to measures of physical well-being and other relevant factors in persons over 74 years need to be explored.

The reasons for excluding infants under 6 months of age from NHANES are unclear. However, there seems to be a need for updating the baseline dietary data on a national sample for this age group, including the extent of breastfeeding and feeding of milk and other foods and supplements.

"Usual" intakes

Consensus on an operational definition of "usual" dietary intake by an individual or group has not yet been reached. In practice, a number of researchers use a specified number of days of actual intake spaced over a stipulated time period. From methodological studies recently carried out at the Nutrition Monitoring Division (NMD) of the Human Nutrition Information Service (HNIS) at the U.S. Department of Agriculture (USDA), there appears to be a limit to the number of replicated dietary reports many people in a survey are willing to provide. In one study of data collection methods, an initial in-person interview was followed by a telephone interview to obtain a 1-day recall in each subsequent quarter. This study resulted in a borderline or, in the eyes of some researchers, an unsatisfactory final response rate. This brings up the subject of response burden.

Respondent burden

The difficulty in collecting multiple days of dietary reports from an individual leads to consideration of ways to lighten the burden or ease the demands on the respondent. In the previous discussion of systematic bias, respondent reporting is mentioned as a likely source of bias in several instances. However, some of the blame for poor quality data may be the fault of survey designers and planners, who are asking for information that is beyond the respondent's ability to provide. Forcing a respondent to give an answer is not solving the problem. Some of the burden in obtaining better quality information could be transferred to more highly skilled and trained interviewers provided with better techniques, questions, and probes eliciting information in a comfortable, nonthreatening setting. This may require much more testing to improve the various elements of the dietary interview. Determining who is the best source of information is important. For example, packaging and brand information might best be asked of the household respondent in the home interview. Also, could the differences in alcohol consumption by males in NHANES and NFCS be a result of males' answering for themselves in NHANES, whereas in NFCS the household respondent may be the source of information but be unaware of the male's total alcohol consumption?

Alternative and multiple methods with possible validation study

The NHANES protocol provides for a home visit about 2–4 weeks prior to the visit to the mobile examination center. This would seem an opportune time to administer a 1-day dietary recall somewhat similar in method to that used in NFCS. Use of model glasses, mugs, cups, bowls, and spoons would not be necessary, as those used by the sample person would be available for actual measurement of capacity or of customary amounts consumed. This might eliminate some of the guessing that is required in relating to dishes, glasses, or spoons foreign to the respondent. Results from the home interview (using a minimum set of aids such as standard measuring cups and spoons, a ruler, and some mounds and shapes) could be compared with results using the full NHANES III set of models to determine validity of results when the full set is used in the clinic setting. I do not recall such a validation study having been made with a suitably large sample. If conducting the dietary interview in the home is not considered feasible, the usual measures of portions eaten in the home cups, glasses, bowls, and spoons could be obtained during the home interview to validate the model measurements indicated during a clinic interview.

Added to these advantages of home interviewing would be the opportunity to obtain multiple dietary reports (including weekend days) and dietary reports more comparable to those from NFCS. It might be possible to conduct an experiment using the NHANES setting and an NHANES interviewer with the NFCS method versus an NFCS home location and an NFCS interviewer with the NHANES method. This might reveal whether methodological differences now considered to be trivial are more serious.

Meal and snack patterning

Meal and snack patterning of combinations of foods eaten together, and of foods eaten in a day and across days, has long been an area of interest and analytical work at the NMD of HNIS at USDA. Some methodological studies have been reported, but more research is certainly needed. Nutrition education, guidance materials, and USDA food plans are described in terms of foods in meals and snacks or budgets. Market analysts and the food industry need information about uses of foods. Increased numbers of ethnic groups have brought their food customs with them; how they adapt or select from among the U.S. foods available is important for their health and important to know in order to provide guidance for planning nutritious meals for their families. Knowledge of low-income and elderly persons' meal patterns as well as those of other subpopulation groups is important for understanding how to deal with their particular problems. Regional differences in dietary practices also must be taken into account.

Internal validity checks

In addition to the methods for checking the internal validity of estimates in a survey previously mentioned, another method is being considered by the NMD. This is a question on frequency of use of calcium-rich foods over a period in the recent past to determine whether designation of frequent and nonfrequent users in the frequency question is similar to that obtained from the NFCS 3-consecutive-day method.

Another internal validity check—a quantified frequency question concerning alcoholic beverage consumption—could serve a dual purpose. If it were included in both NHANES III and NFCS 1987, it would serve as a linkage question that could help unravel why alcoholic beverage consumption reports are higher in NHANES than in NFCS.

Recommendations

- Retain the 24-hour recall in the mobile examination center for comparability and the study of trends and patterns of changes.
- Try to collect multiple days of data (minimum of two) using either an in-home interview or a telephone recall. (Pretesting is needed.)
- Include the frequency question on alcoholic beverage consumption as an internal check and for linkage between surveys.

Second discussant's remarks about State and local perspectives

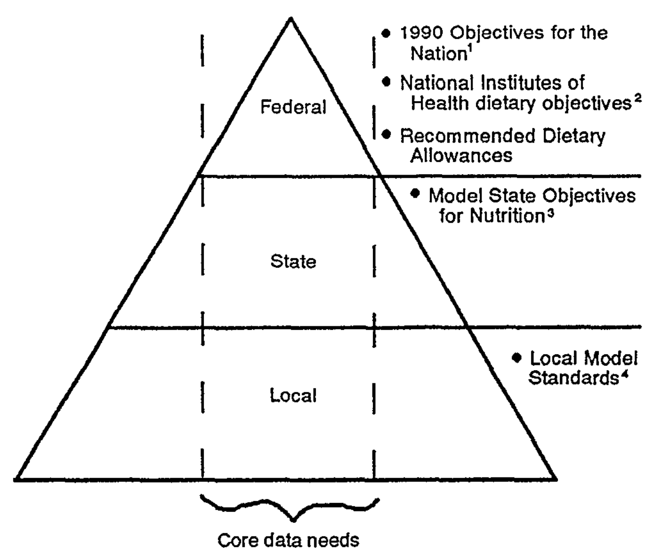
by Susan B. Foerster, M.P.H., R.D., Association of State and Territorial Public Health Nutrition Directors

On behalf of the Association of State and Territorial Public Health and Nutrition Directors (ASTPHND), I would like to express our appreciation for the opportunity to participate in this very stimulating and important working conference.

First, I would like to define what ASTPHND is and what it represents relative to these proceedings. Unlike most persons here, we are not researchers; rather, we are planners of intervention programs, implementors, and evaluators. With one of our most important roles being that of identifying unmet needs in our States, we are "consumers" of the project: National Health and Nutrition Examination Survey (NHANES) III research information.

Second, we are historically "new" consumers. The original purpose of NHANES was to provide the intellectual data with which to set national health and nutrition policies. With the advent of the "New Federalism," including block grants, relaxed performance requirements, and minimal programmatic oversight as well as decreased Federal funds for domestic programs, the Federal Government no longer uses its intervention arm as it did in the past. Rather, it is left to States and localities to establish policies and priorities, to target funds and high-risk groups, and to compete with other sectors of government for public dollars. In other words, the intellectual data base needed for policy setting and competition for sources has now been separated from the levels of government with the capacity and acute need to use the information.

For these reasons, ASTPHND has identified as its top priority the need to develop and strengthen the data and epidemiologic underpinnings available for use by States and localities. NHANES is one of the data bases central to this priority.



[1] Blood pressure; pregnancy and infant health; improved nutrition.
[2] Reduced fat, cholesterol, sodium, alcohol; increased fiber, foods high in vitamins A and C, cruciferous vegetables, and calcium.
[3] Maternal and child health; adult health; environmental nutrition and consumer protection; and administration and finance.
[4] Chronic disease; aging; dental health; maternal and child health; primary care; school health; others.

Figure 1. Public health priorities

State nutrition priorities

Throughout this workshop, there has been extensive discussion about hypotheses. Let me share with you, visually in figure 1, what we believe the "hypotheses" to be. At the Federal level, hypotheses are proposed by the 1990 Objectives for the Nation (1), the various dietary targets set recently by the National Institutes of Health (blood cholesterol, osteoporosis, hypertension, cancer) (2–5), and, of course, the Recommended Dietary Allowances (RDA's) (6). At the State level, States set their own priorities, but ASTPHND is now working on a consensus recommendation, Model State Objectives for Nutrition, which will be available in May 1986 (7). At the local level, the document *Model Standards: A Guide for Community Preventive Health Services* has just undergone its second revision and presents a roadmap of priorities agreed upon by six major bodies of public health policymakers (8). Thus, for us, the "hypotheses" are fairly well established.

Let us look for a moment at dietary hypotheses. In the new Model State Objectives, we will probably adopt the recommendations for nutrient intake in the RDA's as well as the important new food and nutrient recommendations of the National Institutes of Health, including the National Cancer Institute (2–5). The list and groups will, therefore, include fiber, vitamin A- and C-rich fruits and vegetables, cruciferous vegetables, calcium, fat, dietary cholesterol, sodium, alcohol, and certain food practices, such as choosing specific low-fat or high-fiber foods, low-fat food preparation methods, and avoiding certain foods, such as those with naturally occurring carcinogens. Note here that we begin to need information on awareness, knowledge, attitudes, and skills as well as actual food practices. Parenthetically, note that as presently written, neither the 1990 Objectives nor the Local Model Standards contain adequate dietary objectives, although they are somewhat stronger in setting nutrition education targets.

Roles appropriate for NHANES and emerging issues

Throughout the workshop, there has been discussion about what is or is not an appropriate role for NHANES with regard to national data needs. From a State perspective, I would recommend the following.

Nutrient adequacy of diets, especially among children, the elderly, and the poor, will remain a priority. There is no other data system that can provide normative data on diet related to health status.

Health promotion, that is, what individuals can do to keep themselves healthy, is a vital new strategy in public health. We must have a data base that connects health status with the prerequisite education-related population characteristics of awareness, knowledge, attitude, skill, behavior, and service utilization. As there is a real need to describe the relationship between these characteristics and services, NHANES could make a very important contribution toward pinpointing unmet service needs.

Environmental nutrition and consumer protection are traditional public health program areas in which new dimensions are being defined for public health nutrition to address. The growing presence of agricultural and industrial contamination in the food supply is making it increasingly important to know the diversity of the food supply and who consumers of specific foods or types of foods are. One example would be consumers of "unregulated" foods that do not enter the commercial market (for example, wild game and home-grown foods) and the amounts of these foods such individuals are likely to consume. It is upon such estimates that States can use established tolerance levels, assess risk, and determine whether public health action is needed. Besides establishing population norms, NHANES can assume the critical role of providing methodologic expertise to States that need to have suitable survey techniques for handling such problems within their own borders.

Consumer protection includes nutrition labeling of foods, health claims for foods, and nutrition fraud. NHANES is the only data base that can connect these characteristics, not only with dietary intakes—the Nationwide Food Consumption Survey of the U.S. Department of Agriculture (USDA) can do this—but also with nutritional and health or disease status.

Problems of short-term food availability, hunger, emerging malnutrition, and episodic biologic or toxic contamination require rapid response and are therefore not appropriate concerns of NHANES as it now operates. Other surveys' methods and data bases are necessary to address these matters.

NHANES as a longitudinal study

An objective of the workshop is to define data elements needed for retrospective and prospective national population study. Let me preface my remarks by saying that I had been unaware that NHANES is considering itself longitudinal, and it would seem that shifting from a cross-sectional survey has vital policy implications, as follows:

1. Practically speaking, finding ongoing funds is far harder for long-term studies than for those with more immediate payoff. One example with which I am familiar is the Alameda County Human Population Laboratory. Although the lab does superb and often unique work, it spends an extremely high percentage of its time securing ongoing funds. Further, during periods of budget reduction, long-term investment in studies for which the payoff is likely to be years in the future is not likely to be successful. Rather, studies that provide short-term results are likely to be viable. Study results are usually needed yesterday!
2. Historical objectives and unique characteristics of NHANES have included its being a true national cross-sectional snapshot of diet and health, its ability to show national trends over time, and its collection of raw or totally new dietary and health data. If NHANES becomes longitudinal, how will the sample be refreshed so that a true national cross section is maintained in the future? I note that from February to March 1986, the projected sample size of NHANES III was reduced from 60,000 to 40,000. How can 40,000 cover the former participants and new ones? Second, does not the mere exposure to the NHANES examination introduce participant bias in the future examinations, thereby jeopardizing the principal objectives of NHANES related to monitoring national population characteristics? Third, NHANES provides a major source of new or raw information. If for the sake of consistency with previous surveys, the same questions and methods must always be used, does that not erode NHANES' ability to meet changing needs for information?
3. NHANES' contribution to science as a longitudinal study needs examination. Some have stated that a longitudinal NHANES would be superior to clinical trials such as the Multiple Risk Factor Intervention Trial or the Lipid Research Clinics in unraveling multifactorial relationships between diet and health. How likely is it that a study with totally uncontrolled "interventions," such as secular trends, only a fraction of which would be captured in NHANES, would be better at elucidating disease causality than would cohort studies? Rather, NHANES retains its unique value in characterizing and measuring population trends and establishing norms that aid in interpreting findings of clinical trials.

In conclusion, I would urge that perspective be maintained in shifting from a cross-sectional survey. Right now NHANES occupies the unique and absolutely essential position of being the only survey that links health status with an array of other characteristics that States as interventionists are required to know in order to do their job toward improving the public's health. A natural tension exists between the data base demanded by academicians before they recommend action and that expected by public health policymakers; this tension has been well described recently by Miller and Stephenson (9). NHANES is the data base of public health, and it must retain its flexibility and innovative capacities to continue serving its principal users by providing cross-sectional, contemporary, and descriptive population statistics.

States' needs for nutrition-related data

Now let me provide an overview of how States as program planners view nutrition data needs. As depicted in figure 2, achieving optimal nutritional status in the population is a chain of sequential events with multiple intervention points, and data are needed at each step. Consecutively, we might list them as follows. Community, information, food, and social environments support and reinforce the desired health behaviors; these environments provide activities and services that introduce and promote the desired health behaviors. These services lead to accomplishment of educational prerequisites that enable and motivate the individual to choose and practice the desired behaviors. Environment, service, and education together result in the desired dietary behaviors, including factors of preferential food choices, actual selection among various food options, the multiple stages of food preparation, and at-table behavior, such as adding seasonings or trimming fats. These new behaviors lead to the reduction of nutritional risk factors and improvements of nutritional status. In the short term disease incidence, and later disease prevalence, will be modified. Also decreased will be morbidity, including the cost of associated health services and eventually premature mortality. In program planning terms, activities related to intervening in the environment would be called "structure" objectives; activities and services are called "process" objectives; and educational, dietary, and health factors are called "outcome" objectives.

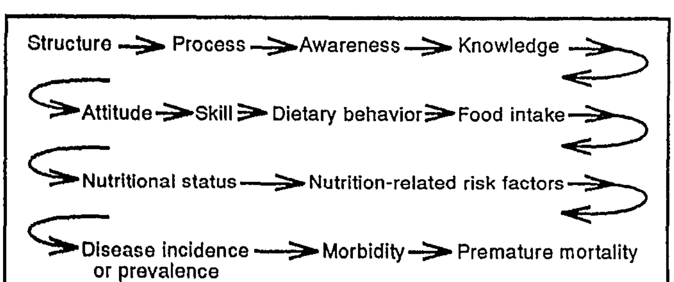


Figure 2. Chain of events influencing diet, nutrition, and health

Table 1. Inventory of current and potential data sources useful in nutrition monitoring

1.	Health and Nutrition Examination Survey
2.	Health Interview Survey
3.	Ambulatory Medical Care Survey
4.	Hospital Discharge Survey
5.	Birth records
6.	National Death Index
7.	Nationwide Food Consumption Survey
8.	Continuing Survey of Food Intakes by Individuals
9.	National Nutrient Data Bank
10.	Food and Drug Administration Total Diet Study
11.	Multi-Purpose Consumer Survey
12.	Biennial Food Product and Labeling Survey
13.	Dietary Supplement Survey State-National Data Bases
14.	Center for Chronic Disease and Health Promotion, Centers for Disease Control School Health Survey
15.	Center for Chronic Disease and Health Promotion, Centers for Disease Control Prenatal Surveillance
16.	Center for Chronic Disease and Health Promotion, Centers for Disease Control Pediatric Surveillance
17.	Center for Chronic Disease and Health Promotion, Centers for Disease Control Telephone Behavioral Risk Factor Surveillance System
18.	Multiple Risk Factor Intervention Trial/Lipid Research Clinics
19.	Hypertension Detection, Treatment and Follow-Up Program
20.	National Public Health Reporting System Categorical Program Data Sets
21.	Women, Infants, and Children Program
22.	Perinatal Regionalization/High Risk Prenatal
23.	Early and Periodic Screening, Detection and Treatment
24.	Adult Programs
25.	Aging programs, miscellaneous data sets
26.	Death records
27.	Genetic Diseases Reports
28.	National Institutes of Health statistics—blood pressure, cholesterol, diabetes
29.	Hospital discharge reports
30.	State registries—tumor, blood pressure, diabetes
31.	Ross Labs
32.	National Institutes of Health Community Trials—Stanford, Minnesota, Pawtucket

SOURCE: Adapted from data sources cited in: Public Health Service. Promoting health/preventing disease: Public Health Service implementation plans for attaining the Objectives for the Nation. Public Health Rep Sept.–Oct. suppl., 1983.

Viewing all available sources that may provide necessary information about these objectives, we can list data bases that have been cited in the 1990 Nutrition Objectives for the Nation (1) as components of the National Nutrition Monitoring System (10) and add to them other data bases that are commonly available in States and localities. This list appears in table 1.

Scattergrams are popular in this workshop; let us look at the question of data categories this way. By comparing the contents of figure 2 with those of table 1, we can identify clusters of relative data strength with other areas of data weakness. Thus, table 2 reveals clearly that major categories of data collection that need to be strengthened include structural, process, and educational parameters. Inasmuch as NHANES is the only national data source that can link these parameters with health status, it becomes imperative for NHANES III to add more data elements to these fields.

Vertical integration and use of data

Implicit in this discussion is the assumption that data collected in NHANES are to be selected with consideration to their utility to States and localities. Thus, the

Table 2. Sources for and gaps in nutrition data needed by States and localities for policy, planning, and evaluation

Data collecting agency	*Structure*	*Process*	*Education*	*Dietary*	*Health*
			Data base		
National Center for Health Statistics. . . .	1–5	2,3,5	1–5	1	1–4,6
U.S. Department of Agriculture	1,2	---	7–9	7–9	---
Food and Drug Administration	10–12	---	10,11	10,13	---
Centers for Disease Control. . . .	14	14	---	---	15–17
National Institutes of Health	---	---	---	18,19	18,19
State categorical programs.	20	20–25	---	---	15–17,26–30
Other, miscellaneous, and proprietary	31	32	31,32	---	---

selection of actual data elements that are consistent with those already in use or selected as priority for collection is a fundamental recommendation from the States. It is, therefore, strongly recommended that NHANES avail itself of the work now underway within ASTPHND to identify a "national core data set for nutrition," which is based in large part on existing 1990 Objectives and Local Model Standards but also encompasses new parameters, such as those called for by the National Public Health Reporting System and those that are needed for new national initiatives, such as the National Cancer Institute's Year 2000 Dietary Objectives (5,11). The deadline within ASTPHND for completion of a preliminary "core data set" is fall 1986. Compatibility among these three levels of users would immeasurably maximize the usefulness of NHANES III in public health.

Optimizing NHANES III

A central question posed at the workshop has been that of how to increase the value of the NHANES III investment. From a State and local perspective, a "macroview" of NHANES, the following specific recommendations seem to be the most important.

Select data elements that are vertically integrated, as discussed above.

Maintain comparability of methods to assure comparability of results. For example, the information that the American population has been consuming about 37 percent of its total calories from fat since NHANES I is, for those of us who are designing intervention programs, vital and surprising information. Such statistics not only dictate the content of the programs we design but also dispel the perception that nutrient intake is improving through secular trends alone. If, indeed, major reduction in heart disease and cancer are likely at 30-percent fat levels, then it is clear that public health authorities must reassess the priority they place on nutrition as a prevention strategy.

Lateral interdigitation with other Federal data bases is necessary in order to fully utilize each survey, avoid duplication, and prevent incompatibilities that introduce "noise" rather than clarity into policy questions. For all the agencies that are part of national nutrition monitoring, linked data elements need to be identified that correspond to the priority data element selected by State and local users. For example, USDA and the National Center for Health Statistics ought to be able to agree on 24-hour dietary recall methods and other collection of raw dietary data within the Nationwide Food Consumption Survey that would, in turn, generate an updated and periodically revitalized List-Based Food Frequency for use in NHANES.

Sample design should consider the needs of non-Federal users. Drawing a sample that allows only national projections fails to meet the needs of the vast majority of users. Major target groups such as Hispanics, Asian-Pacific Islanders, Native Americans, and the very old need to be added to the traditional population segments. Perhaps more important is sample design to permit subnational estimations according to regions that make sense in terms of health or lifestyle rather than regions selected arbitrarily by the Department of Health and Human Services, USDA, or the Department of Commerce. Imagine the value of NHANES reports labeled the "New England" region or "Pacific Northwest" region! In addition, it would seem reasonable to design NHANES interviews and analyses such that States or regions could add participants to the sample or optional data elements to the standard examination, even if this were done on a cost-reimbursable basis. Not only would this add to the subnational value of the data, but also more localized "ownership" of NHANES would result.

Technology transfer is a subject that has been little mentioned. Survey methods selected for NHANES are scientifically superior and technologically advanced. Therefore, States and localities look to NHANES for prototype approaches and methodologic guidance. Of course, it is vital to continually include new methods, such as the collection of "matter," as proposed earlier in the workshop. I am surprised that no one has yet mentioned the inclusion of "functional parameters" of nutritional assessment, especially for children or the elderly. Tests of cognition, manual dexterity, and memory related to nutritional status would seem to be farsighted new dimensions for NHANES. However, it is of greater importance to consider the potential for broad public health application in the development of survey methods. Total automation of dietary interviews, particularly if suitable for microcomputers, and shortened questionnaires or validated telephone followup interviews would be extremely valuable to public health practitioners. In addition, provision should be planned to speed the dissemination of new technology within the public health and research communities.

Finally, attention should be paid in advance to a *data delivery system* for NHANES. In the Hispanic HANES major strides have been made to release data promptly and into the hands of university and other non-Federal users. Although this represents major progress over the lengthy delays of the past, it could be improved by construction of a *planned program of data analysis* in partnership with non-Federal experts to assure that major policy questions are investigated. Otherwise, such examinations are serendipitous, and there are well-known instances of data misuse, unauthorized publication, and subsequent public confusion. In addition, the release of *intermittent and timely reports of findings* would greatly enhance the usability of NHANES. As has been done in Hispanic HANES, it would seem desirable to convene an advisory mechanism to aid the NHANES staff in setting ongoing priorities for data analysis and in disseminating information about NHANES.

It has been my pleasure to participate in this workshop representing State perspectives. We shall watch with acute interest the progress of NHANES III and look forward to working together toward achieving mutual data objectives.

References

1. U.S. Department of Health and Human Services. Promoting health/preventing disease: Objectives for the Nation. Washington: U.S. Government Printing Office. 1980.
2. National Institutes of Health. Lowering blood cholesterol to prevent heart disease. Consensus Development Conference Statement 5(7). Bethesda, Maryland. 1984.
3. National Institutes of Health. Osteoporosis: Cause, treatment, prevention. NIH Publication No. 86–2226. Bethesda, Maryland. 1986.
4. Subcommittee on Nonpharmacological Therapy, 1984 Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure. Nonpharmacological approaches to the control of high blood pressure. Hypertension 8(5):444–67. 1986.
5. National Cancer Institute. Cancer control objectives for the Nation: 1985–2000. NCI Monographs 2. Bethesda, Maryland. 1986.
6. National Academy of Sciences, Subcommittee on the Tenth Edition of the Recommended Dietary Allowances. Recommended Dietary Allowances. 10th rev. ed. Washington: National Academy Press. 1989.
7. Foerster SB, Lee S, Bonam S, et al. Nutrition goals for three levels of government: Recommended model objectives for States [Abstract]. Paper presented at the American Public Health Association, Oct. 1, 1986.
8. Model Standards Work Group: Model standards: A guide for community preventive health services. 2d ed. American Public Health Association, Association of State and Territorial Health Officials, National Association of County Health Officials, U.S. Conference of Local Health Officers, and the Centers for Disease Control. 1985.
9. Miller SA, Stephenson MG. Scientific and public health rationale for the Dietary Guidelines for Americans. Am J Clin Nutr 42:739–45. 1985.
10. Interagency Committee on Nutrition Monitoring. Nutrition monitoring in the United States—The directory of Federal nutrition monitoring activities. DHHS pub. no. (PHS) 89–1255–1. Washington: Public Health Service. 1989.
11. U.S. Department of Health and Human Services. Healthy people 2000: National health promotion and disease prevention objectives. DHHS pub. no. (PHS) 91–50212. Washington: Public Health Service. 1991.

Chapter 7
Working group consensus statements

Working Group Number 1: Statistical issues related to the design of dietary survey methodology for NHANES III

Kiang Liu, Ph.D., Northwestern University, Harold A. Kahn, M.A., consultant in epidemiology, and George H. Beaton, Ph.D., University of Toronto

The recommended methodologies are as follows.

1. A 24-hour recall should be collected from each individual to maintain comparability with the Nationwide Food Consumption Survey.
2. Some type of food frequency data should also be collected.

Concerning multiple 24-hour recalls, several points should be made.

1. Multiple 24-hour recalls should be collected, preferably from each individual. If that is not possible, multiple 24-hour recalls should be collected in a randomly selected subsample.
2. Professors Beaton and Kahn felt that a minimum of two 24-hour recalls should be collected either from each individual or in a randomly selected subsample. Dr. Liu felt that the minimum number of replicate recalls should be three.
3. Multiple recalls of the number being recommended will permit the adjustment of distributions for within-person variability. This procedure, as outlined in *Nutrient Adequacy: Assessment Using Food Consumption Surveys*,[a] would permit more accurate estimates of prevalence. Dr. Liu felt that the correction of distributions based on two 24-hour recalls should be undertaken with great caution. He was not confident that a good enough estimate of the ratio of within-person to between-person variability could be obtained with two replicates; he would have more confidence in an estimate of the ratio that was based on three replicate 24-hour recalls.
4. Although multiple recalls will allow for more accurate estimates of distributions, the number of replicate recalls recommended will not permit the identification of particular individuals who are above or below a particular level of nutrient intake.
5. For comparisons between a dietary variable with a significant nonzero ratio of within-person variability to between-person variability and a nondietary variable with a zero or essentially zero ratio of within-person to between-person variability:

   a. When the dietary variable is the dependent variable, variability is controllable by sample size.
   b. When the dietary variable is the classification (independent) variable, the key issue is the misclassification error associated with the dietary variable's large ratio of within- to between-person variability. It should be recognized that the number of replicate measures recommended may result in only minimal enhancement of the ability to do hypothesis testing between or among the dietary variable's classification groups.

Two points are made concerning validation of dietary methods:

1. Professor Kahn recommended that the 24-hour recall be validated for estimating group means. This would be accomplished by comparing mean urinary excretion of a few selected nutrients (such as nitrogen excretion for protein) with mean intake assessed by the 24-hour recall. The study is outlined in his discussant's remarks on Dr. Liu's paper.
2. Drs. Beaton and Liu felt that any study designed to assess the validity of the 24-hour recall should be undertaken to estimate intake for individuals rather than groups. Professor Kahn does not disagree with this, but he knows of no natural method for validating individual intake in a free-living population.

---

[a]National Academy of Sciences, National Research Council, Subcommittee on Criteria for Dietary Evaluation, Coordinating Committee on Evaluation of Food Consumption Surveys. Nutrient adequacy: Assessment using food consumption surveys. Washington: National Academy Press. 1986.

Working Group Number 2: Recommended dietary methods for cancer

Gladys Block, Ph.D., National Cancer Institute, Walter C. Willett, M.D., Dr.P.H., Harvard University, and Gail McKeown-Eyssen, Ph.D., Ludwig Institute for Cancer Research

NHANES III could provide potentially important data for the prospective analysis of the relationship between diet and risk of cancer, as well as for diet and other diseases, if appropriate dietary methods are used. Although larger cohorts exist, none provides the opportunity, as NHANES III could, to examine the role of diet in adolescence or early adulthood in the development of disease many years later. In addition, cross-sectional studies on the relationship of baseline diet to biochemical parameters would be useful. NHANES III is not a data set in which it would be desirable to attempt case-control studies involving the collection of retrospective dietary data for all respondents at baseline, although future studies on a subset of this population might shed useful light on whether observed age differences in intake reflect cohort differences or changes associated with aging.

In order to make prospective approaches fruitful, long-term followup and reassessment at 6- to 10-year intervals would be required. Potential confounding variables must also be assessed at baseline and reassessed at those intervals, in addition to the collection of the dietary data. The collection and storage of biological samples for the subsequent analysis of nutritional-biochemical factors and disease and the collection of urine and stool as potential indicators of actual dietary intake could also provide important data.

The usefulness of NHANES III data for the prospective or cross-sectional analysis of the relationship between diet and disease or between diet and biochemical parameters is predicated on the use of a dietary method that is valid for the assessment of an individual's usual diet. Data from 2 or 3 days, or from the sort of frequency questionnaire used in previous NHANES surveys, are not adequate for this purpose, nor will they permit substantial advancements in the level of our understanding of the relationship between diet and disease or diet and biochemical parameters. For many of the nutrients that have been examined, the data are quite clear that extensive quantified frequency and/or list-based diet history methods provide a representation of an individual's usual diet at a level of precision equal to or better than that provided by 7 or more days of recall or records. Furthermore, a combined frequency-history method will provide the opportunity to examine disease relationships at the level of foods and food groups as well as at the level of nutrient intake.

An extensive frequency-history method is defined as containing 75–100 items and taking approximately 25 minutes for completion. As list-based instruments are relatively simple to complete and generally do not require equipment (such as food models), assessments of changes in diet over time can be made by readministering the questionnaire by mail. To this should be added some brief questions on food preparation, water consumption, and so forth. Information on vitamin supplement use is essential.

Unless NHANES III uses an extensive frequency-history or many (7–21) multiple days of recalls, analyses of the survey will be limited to descriptive studies of population characteristics and will not allow analytic study of diet-disease relationships at the individual level because of the substantial misclassification and the attendant loss of statistical power. Descriptive studies are the least powerful way of examining etiologic relationships and the one most subject to hidden and "ecologic" confounding. Analytic studies require data valid for an individual's usual diet.

Data from 24-hour recalls can provide internal validation, accurate population means, and comparability with prior surveys (although note should be taken of the existence of such data from other national surveys). Replicate 24-hour recalls should be performed on randomly assigned days, including weekends, approximately 6 months after the first measure to eliminate seasonal bias.

Therefore, it is recommended that NHANES III use an extensive 75–100 item frequency-history as the individual measure of nutrient intake, in conjunction with a single 24-hour recall on every participant and replicate measures on either everyone or a subset. If respondent burden or fiscal considerations require a choice between the two methods, it is recommended that NHANES III use the extensive frequency-history, if the goal is to use the health and biologic data for meaningful analytic studies of diet and disease.

Working Group Number 3: Recommended dietary methods for cardiovascular disease

Patricia J. Elmer, Ph.D., R.D., University of Minnesota, Arlene Caggiula, Ph.D., R.D., University of Pittsburgh, and Nancy Ernst, M.S., R.D., National Heart, Lung, and Blood Institute

As NHANES moves from a cross-sectional design to a longitudinal design (for research and hypothesis testing), the hypotheses and goals for future data collection need to be clearly identified and the timeframe for evaluation stated. Dietary intake information is extremely important for questions related to the etiology of cardiovascular disease (CVD), particularly as NHANES III data will serve as a baseline for longitudinal studies.

Specific dietary assessment tool

Questions arise about the use of multiple 24-hour recalls versus food frequency lists in relation to solving the problems of intraindividual variance and misclassification and preserving some information about specific foods. For cardiovascular disease, the dietary assessment tool needs to obtain information about nutrients associated with blood pressure regulation and blood cholesterol levels. These nutrients include sodium, potassium, and specific fatty acids. There is concern that the use of a food frequency list could result in misclassification of individuals at high or low intakes of some nutrients of interest because of systematic biases, in particular those related to the construction of the instrument, but also the choice of food composition values and even food portion assumptions.

Some of these questions or problems could be alleviated by lengthening the list of the food frequency instrument or using differential portion sizes, but a resulting problem then becomes a long interview and a considerable increase in respondent burden. (It is important to note that in some situations food frequency questionnaires have been used as short, quick dietary assessment methods for epidemiologic studies, and in many cases they can still provide meaningful dietary intake data for groups.)

Nutritional epidemiology

Among CVD nutritional epidemiologists, reservations related to the use of a food frequency questionnaire relate to several specific areas. There is concern that there should be further analyses of data, examining reliability and validity among different populations (age-ethnicity) before this method is used as the only dietary method. Questions to be answered include the following: Is this a valid and reliable method to assess dietary intake of children? Should the food list be different for specific age and ethnic groups? If so, what is the implication of interpretation if different lists are used?

Important unresolved etiological questions related to CVD and hypertension could be addressed in proposed prospective analyses planned for NHANES III. In particular, NHANES III could make a unique contribution from the data obtained by the oversampling of black persons and other minorities, the elderly, and the low-income groups. Important in these etiological evaluations is information on the interaction of nutrients and such variables as body weight and physical activity. NHANES III then may be able to address questions such as whether there are critical ages at which interactions affect specific disease outcomes.

If a longitudinal component is added to NHANES III, the dietary assessment tools must facilitate measurement of dietary change. If the impact of specific health promotion programs (for example, the National Cholesterol Education Program) is of interest, information on dietary attitudes, knowledge, and eating behaviors might also need to be assessed.

In consideration of the above-mentioned factors, information presented in the manuscript by Dr. Patricia Elmer (chapter 3), and discussions at the NHANES III Dietary Survey Methodology Workshop, both the 24-hour recall and the food frequency list are recommended as being suitable for NHANES III. Improvement and modification of the food frequency methodology is further recommended.

Working Group Number 4: Recommended dietary methods for energy balance

Dorothy Blair, Ph.D., Pennsylvania State University, Johanna Dwyer, D.Sc., R.D., New England Medical Center, and Theodore Van Itallie, M.D., Columbia University

There is general agreement that an attempt to measure energy balance is an inappropriate aim for NHANES III. Energy balance is an issue of quantifying energy input and energy expenditure. Balance tends to occur over varying periods of time within individuals, as indicated by peaks and valleys that negate each other when a long-term stable, but short-term fluctuating body weight is plotted over time. Even very precise measurements of energy intake and expenditure would not adequately categorize individuals who achieve energy balance over a longer time than is possible to measure in a cross-sectional survey. Furthermore, such precise measurements are not even within the realm of possibility for NHANES III.

A change in fat stores is the measurable outcome of an energy imbalance and may be more precisely approximated in a survey than either energy intake or energy expenditure. Serial measurements of body fat are suggested if useful information on the degree of energy imbalance in the population is a priority. Such information would be useful for interpreting energy intake data. Weight histories and more specific information on restrained eating behavior would also be useful.

Although energy balance is not considered to be an appropriate aim for NHANES III, precise, accurate, and valid measurements of total usual energy intake and habitual activity levels are appropriate aims. There is concern that possible estimate biases resulting from obesity be recognized and minimized. With regard to dietary assessment, probing for information on snack and dessert food items, increased attention to techniques for eliciting accurate portion size estimates, and increased frequency and variety of days assessed would possibly reduce respondent bias resulting from obesity. To this end, the home interview could be used to collect dietary data, particularly on serving container sizes, portion sizes, and specific brands and snack food items. The food frequency could be used specifically to probe for snack food and alcohol use.

It is recommended that random repeat 24-hour recalls accompanied by a semiquantitative food frequency be considered. Both weekday and weekend dietary intakes should be assessed. Continuing to obtain one 24-hour recall in the mobile examination center would preserve historical comparability. Subsequent telephone-administered 24-hour recalls are a possibility, especially to make weekend data collection more practical. This technique requires further validation, particularly with regard to portion size estimations. Ancillary data that should be collected include use of medications that affect energy intake (such as anorexic agents), cigarette smoking, and other drugs, such as antidepressants, that act as orexigenic agents.

In summary, energy balance cannot be effectively assessed under conditions outside of the laboratory setting (that is, a metabolic ward). Even if energy balance could be precisely measured, its utility would remain highly questionable because fluctuations about the mean may render the concept of energy balance meaningless in the short term.

Working Group Number 5: Recommended dietary methods for osteoporosis

Nancy E. Johnson, Ph.D., University of Hawaii, and Mary Fran Sowers, Ph.D., Cornell University

To characterize the current diet for all respondents, the following recommendations were made:

1. Multiple 24-hour recalls—as many days as feasible (in the range of 2–3); use of telephone interviews is acceptable.
2. A quantitative food frequency that focuses on food sources of nutrients with high intraindividual variability, such as vitamin A (if time allows).
3. Current vitamin-mineral supplement use—characterize in as much detail as possible frequency of use and amount of nutrients in each supplement taken.
4. Current water intake, consumption of water and water-based beverages, use of water softener or charcoal-based water conditioner, and analysis of community water supply for all minerals possible, including calcium, fluoride, lead, and cadmium.

To characterize the past diet for all respondents, the following recommendations were made:

1. Diet: Ask about daily milk consumption separately ("usually" or "always," "daily" versus "never" or "rarely") and daily cheese, yogurt, and ice cream consumption for four age groups—0–5 years, 10–14 years, 15–24 years, and 25–44 years.
2. Vitamin-mineral supplement use: Ask about use of supplements and type (for example, multimineral versus simple nutrient) or obtain as much detail as possible, but not brand information, for 5 years ago, 10 years ago, and 15 years ago.

Working Group Number 6: Recommended dietary methods for nutrition monitoring

Elizabeth A. Yetley, Ph.D., R.D., Food and Drug Administration, Eleanor M. Pao, Ph.D., U.S. Department of Agriculture, and Susan B. Foerster, M.P.H., R.D., Association of State and Territorial Public Health Nutrition Directors

Surveillance of food intakes, nutritional status, and nutrition-related conditions is important for policymaking and program planning. Nutrition monitoring requires the measurement of changes in food and nutrient intake over time; therefore, both baseline and comparable followup data must be available. Estimates of "actual" intakes rather than relative intakes among groups are important. Reasonable estimates of frequency distributions of intakes are also needed to identify prevalence of marginal or unsafe intakes in population subgroups. NHANES is a cornerstone of the National Nutrition Monitoring System (NNMS) because it contains a series of surveys in which unique data are collected, including biochemical and medical measures of nutritional and health status as well as data on dietary intakes and vitamin-mineral supplement use. Another cornerstone of NNMS is the Nationwide Food Consumption Survey (NFCS).

To meet the measurement needs of a monitoring system, methodologies among past, current, and future surveys must be consistent or comparable as well as documented so that trends in food intake and dietary status can be attributed to real changes in the population and not to changes in methods or data bases. In addition, some means of correcting or adjusting for measurement error resulting from intraindividual variation in daily intakes is needed in order to obtain population frequency distributions. To facilitate comparability with past NHANES and with NFCS, and for nutrition monitoring purposes, the following are recommended:

- Continued use of the 24-hour recall.
- Two 1-day independent recalls for all NHANES III participants.
- Alternative approaches to the latter could involve obtaining multiple intake days on a subsample, obtaining subsequent 24-hour recalls by a telephone interview, or adjusting the distribution of nutrient intakes using calculations of intraindividual reliability obtained from NFCS.

Use of the 24-hour recall in NHANES III would provide consistency with previous data as well as group estimates of nutrient intake; however, it is noted that the methodology will be different depending on how and where the 24-hour recalls are collected.

Additionally, linkage and comparability between NHANES and NFCS also are important for NNMS. NHANES III and NFCS should aim to develop common definitions for key variables and to use the same core food composition data base and coding system. Whenever possible, the interviewing techniques, level of probing, and interviewer training also should be consistent. An automated, interactive computerized system to be used in the interview process is strongly recommended. Increased use of the household interview is also recommended as a way to obtain supplementary dietary information. Quantitative information on vitamin-mineral supplement use, water intake, medication use, and alcohol consumption is important for the estimation of complete nutrient or food component intake (as differentiated from food intake).

In addition, information on meal patterns (foods typically eaten together) provides useful data on interactions among food components and on consumer food use behaviors. A focused questionnaire to obtain information on infrequently eaten foods that are important, nutritionally or toxicologically (organ meats, seasonal fruits and vegetables, and so forth), is also valuable.

Chapter 8
Dietary methodology considerations for NHANES III

by Lenore Kohlmeier, Ph.D., Department of Epidemiology of Health Risks, Federal Health Office, Federal Republic of Germany

As in any nutrition survey, selection of the proper dietary assessment tool for NHANES (National Health and Nutrition Examination Survey) III depends on the study goals, resources, and limitations. A number of potential goals for NHANES were presented, argued, and reevaluated during the workshop. The ability of various widely utilized dietary methods to meet these goals partially or completely was discussed. The accuracy, precision, and validity of these methods to measure that which was desired (mostly typical eating behavior) were the overriding topic of deliberations. The need for information on dietary behavior in the study of the etiology of specific diseases resulted in conclusions for specific diseases ranging from realization that the hypotheses cannot be tested in this study to the expectation that acute and long-term trends will be examinable within this study design. Finally, the practical aspects of dietary methodology were discussed, including the need for automation of information collection and calculation and for further development along these lines, even though new methods, which will need validation, are required.

Study goals

The scientists utilizing NHANES III information expect: (1) to assess cross-sectionally diet and risk for disease interactions; (2) to look at the effects of prior behavior on current disease or risk; (3) to look at the impact of current dietary behavior on future disease states through longitudinal sampling; (4) to evaluate the effect of change in the behavior on change in risk of disease through longitudinal sampling. It should be mentioned that not all foods, nutrients, or dimensions of dietary behavior that will be important to future epidemiological hypothesis-exploration are identifiable in advance.

Other purposes, such as those described by Woteki (appendix VIII) in the area of monitoring dietary behavior of the population, are not exclusively NHANES tasks. Each of the four above-mentioned purposes of NHANES needs to be considered in the selection or rejection of dietary methodologies for NHANES III. The responsibility for currently limiting the future resource data base that will result from NHANES III is considerable. Decisions about dietary methodology need to be made in consideration of the strengths and weaknesses of individual methods and the study design. Full knowledge of the consequences of selection or rejection of potential methods is necessary, and adequate cost-benefit argumentation needs to underline the final decision. In addition, characteristics of individual methods can be fatal flaws for certain assessment purposes and must therefore be critically examined in the context of each study goal.

Fatal flaws

Fatal flaws are limiting characteristics of a particular dietary method, making it inappropriate for one or more of the goals of information collection. Wrong timeframes, response bias, invalidity of a method, misclassification of individuals or groups, population compression (or its opposite), high interindividual variation, atypical behavior, and expense may be fatal flaws of a particular method in a study incorporating dietary assessment. Some areas of study may concern themselves with current eating behavior, others are interested in projection of future eating behavior, and still others are acutely interested in behavior that took place in the distant past. Although projections are often made assuming that current behavior is reflective of prior behavior, this may be a misassumption and result in inappropriate and inaccurate conclusions.

If a method results in response bias, with the subjects reporting their dietary behavior differently because of the method of questioning, or different individuals within a population responding differently to the same question, the result may be unusable for the comparisons of population groups that are desired. There are numerous aspects of validity, one of the most important being whether the method measures what it sets out to. For example, 24-hour dietary recalls set out to collect information on the typical short-term cross-sectional eating behavior of a group. Recalls provide valid results only if the individuals can accurately report what was consumed the previous day and if the day selected represents a period of usual behavior. It is invalid when quantity information is required, and it is false when inappropriate numbers of the population have eaten atypically, for example, because they knew that the next day they would have an examination.

Identification of groups at risk based on quantifying dietary behavior depends on accurate placement of the individual within a population distribution. The measurement desired is the usual intake of certain foods or nutrients. If the persons who, for example, eat the smallest quantities or are the least frequent consumers of meat wind up in a higher percentile of the population because of the type of questioning, analyses of risks will be false.

Misclassification becomes less when fewer groups are used for categorization. Compressing the breadth and diversity of consumption into a few broad categories can, however, result in a vast majority of individuals falling into one rubric. For example, asking about the frequency of consumption of a particular food in the categories of many times a day, daily, several times weekly, several times monthly, or rarely may cause very uneven distributions for foods such as coffee, milk, or certain types of fish. The danger is that the vast majority of the population falls into the very high frequency group in the former cases or a very low frequency group in the latter case.

The method of assessment may in itself create atypical dietary behavior and this, in turn, may be a fatal flaw. For example, food diaries, if kept over a series of days, tend to result in reduced consumption (or reporting) of complex meals and recipes. This change in behavior, caused by the methodology, may defeat the purpose of assessment.

If the interest is in information on individual intakes for correlation purposes, selection of a method that measures a very short cross section of dietary behavior may result in intraindividual variations that are as great as or greater than interindividual variations. This purpose may also be a fatal flaw. If the ideal method for a certain purpose is selected but is too costly to be carried out completely in the number of individuals required, this too is a fatal flaw. Fortunately, this fatal flaw can be prevented or corrected.

Dimensions of dietary behavior

This workshop addressed measurement of dietary behavior directly through questioning and indirectly through body fluid measurements and other biological markers. It was agreed that the area of biological markers is not fully developed, and many satisfying solutions of dietary assessment problems may be resolved through future developments. However, for the range of foods and nutrients of interest in NHANES, biological markers will never be exclusive reflectors of dietary behavior; the breadth of interests in nutrients, foods, and processes is too extensive.

Direct questioning was discussed in terms of obtaining information from individuals and groups regarding the quality of their food intake and converting it into daily nutrient intake per person. Dietary nutrient intake is of particular interest to researchers in the field of cancer and cardiovascular disease epidemiology. Not discussed at this workshop were other dimensions of dietary behavior such as food preparation (grilling, baking, boiling, microwaving), methods of preservation (salting, smoking, deep freezing, packaging), the necessity of collecting information on brand names of certain food items consumed, and specific information on the meal itself, such as what was eaten together at a sitting, how frequently meals are consumed, and the meal location (where it was prepared, where it was consumed, and so forth). The extent to which recipes and ingredients in mixed dishes need to be collected was also not discussed. The following list summarizes these factors.

- What was eaten (food groups, recipes, items, brand names)?
- How much was consumed?
- What was discarded?
- How was it prepared and where served?
- How was it previously preserved?
- How was it packaged prior to consumption?
- What was eaten together at the same meal?
- Which snacks were consumed?
- What ingredients and how much comprised "mixed dishes"?
- Were the items consumed "convenience foods" or "homemade"?
- What was the fluid intake, including water?
- What was the current and prior consumption of vitamins or mineral supplements?
- What was the use of "food drugs," such as garlic pills, pollen, protein powders, and lecithin preparations?

Aside from direct measurement of previous eating behavior, the assessment of nutrition-related and health-related histories and dietary knowledge or ignorance were mentioned.

Dietary constituents (nutrients)

Special disease-related interests led to a focus on particular constituents of foods to be quantitatively measured in NHANES III. This list included the basic macronutrients, vitamins, minerals, cholesterol, and fiber. They can be found in table 1.

Table 1. Dietary constituents of interest for particular diseases

Constituent	Disease of interest
Total energy consumption	Cancer, cardiovascular diseases, obesity
Macronutrients	Obesity
Fat	Cancer, cardiovascular diseases
Carbohydrates	Obesity
Protein	Obesity
Alcohol	Obesity, cancer
Vitamins:	
Vitamin A and carotenoids	Cancer
Vitamin C	Cancer
Vitamin D	Osteoporosis
Vitamin E	Cancer
Minerals and elements:	
Calcium	Osteoporosis
Potassium	Cardiovascular diseases
Phosphorus	Osteoporosis
Magnesium	Osteoporosis
Fluoride	Osteoporosis
Selenium	Cancer
Other constituents:	
Fiber	Cancer
Cholesterol	Cardiovascular diseases

Ideal dietary methodology

The following was proposed as a definition of the ideal method for dietary assessment for scientific purposes:

> This ideal method would provide exact and valid information about typical food items and amounts consumed (their preparation and previous preservation) per person prior to the onset of disease. It should be inexpensive, requiring no specially skilled or trained personnel, little or no personal contact, and provide immediate results.

Unfortunately, all experts agreed that it does not yet exist, and compromises and stylization of methods to goals are required. Discussions of methods were limited to those methods with which adequate experience was available. Furthermore, they were based upon the strengths and weaknesses of individual methods, interest study constraints, and fatal flaws.

24-hour recall

Weaknesses of the 24-hour recall included cost of collection and data processing, altered eating behavior on the day before examination, and the large daily intraindividual variation in eating behavior. For individual analyses and subgroup analyses, the recommendation was for multiple 24-hour records—at least two, to allow the estimation of within-person to between-person variability. This is important for individual results as well as for accurate estimation of the population distribution of food or nutrient intakes. A more complete discussion of the statistical issues can be found in chapter 1.

At least a single 24-hour recall is strongly desired to allow NHANES III to continue to be used as a longitudinal measure of national behavior and to ensure some consistent and comparable methodology with the Nationwide Food Consumption Survey.

Food frequencies

Food frequencies were also intensively discussed as an inexpensive approach toward collecting more long-term information on the intake of particular foods, as needed for studying the dietary component of specific diseases (such as cancer and osteoporosis). Although assumptions are generally made that food frequency responses are valid, prepublication data were presented showing misclassification of up to 30 percent of individuals into wrong quintile extremes for a number of foods. Quantitative use of food frequency information was generally not recommended. The validity, usefulness, and comparability of responses among different ages and ethnic groups was questioned.

Food frequency questions have little basic commonality; they differ in the foods or food groups addressed, the range of response possibilities, their inclusion of portion size components, the underlying timeframe, and the way the questions are posed. The right food frequency questionnaire for NHANES III will need to be tailored to the diseases of interest, their etiological-developmental frame, and logistic considerations. For feasibility reasons, an unwillingness to rely solely on 24-hour recalls, and a range of opinions about the usefulness of food frequency questionnaires, it was recommended that some form of food frequency be included in NHANES III. No detailed discussions of what the food frequency should include were undertaken.

Dietary histories

Clear distinctions between what some persons call food frequencies and diet histories were not made at this meeting. The diet history method is generally recognized as desirable for its breadth and quality of information collected, but it is considered too costly in time and resources for the information collected and too cumbersome in the coding, entry, and analyses to be recommended within a large-scale study.

It was noted that the diet history is more widely applied in Europe, and 24-hour recalls, more in the United States. Full automation of the diet history method could make it an attractive alternative to the other methods discussed. Although technically possible, this would require a generation of development and validation. Thus it cannot be recommended for NHANES III, in which a complete methodology should be established and made final in the middle of 1987.

Combinations

More than one method of dietary assessment will be needed in NHANES III to ensure that most of the wide variety of goals are served: Comparability with the past, high-quality data for future longitudinal studies, good quantitative data on individual and group intakes, and clear classification of persons into groups of high or low consumption levels. One 24-hour recall and a food frequency questionnaire represent a minimal program. Multiple 24-hour recalls, at least in subsamples, are strongly recommended by all statisticians for addressing many questions of diet-health interactions. Whether these should be conducted on particular days of the week, in certain seasons, or systematically or randomly requires further discussion.

Automation

An underlying theme of the workshop was the extent to which automation of dietary assessment should and must be undertaken in NHANES III. Automation can take place in a number of individual steps, from the questioning of the subject to the coding of the food information. The entry and checking of codes, the calculations of quantities consumed—taking household measures, losses, conversions, and waste into account—and the conversion of food intake into nutrient intake are all steps that would greatly benefit from automation. Most experience gained in automation has been acquired in the later steps of dietary assessment—the conversion to nutrients. Very few attempts have been made at fully automating dietary assessments, particularly the data capture steps. However, financial strains, positive (albeit limited) experiences with the automation of dietary methods in population studies, and time factors make enhanced automation inevitable. The advantages of full automation are as follows:

- No between-interviewer variability (standardization).
- No response bias introduced by the interviewer (objectivity).
- No change in dietary behavior induced by the method of assessment.
- Scientists decide on exact phrasing, depth of probe, and sequences of questioning.
- No hand coding of food information into machine.
- Built-in and immediate editing and checking (validation) routines (automated quality control).
- Immediate results.
- Better comparability of results.
- Reduced costs of collecting and assessing dietary information.

The most potent of these advantages currently are the elimination of costly and error-prone procedures, the immediacy of results, and reduced costs of operating the study.

Telephone interviews

Longitudinal followup of individuals as well as multiple 24-hour recalls require multiple contacts with subjects. To reduce costs the application of telephone interviewers for either repeat 24-hour recalls or repeat food frequencies entered into the discussions. These would imply new methodologies, requiring prior investigation and validation, but they are considered worth pursuing.

Chapter 9
Process and rationale for selecting dietary methods for NHANES III

by Christopher T. Sempos, Ph.D., Ronette R. Briefel, Dr.P.H., R.D., Clifford Johnson, M.S.P.H., and Catherine E. Woteki, Ph.D., R.D., National Center for Health Statistics

Since the first National Health and Nutrition Examination Survey (NHANES I), which took place in the early 1970's, the National Center for Health Statistics (NCHS) has conducted periodic national nutrition surveys (table 1, (1–6), and appendix VIII). A goal of these surveys has always been the production of national estimates of the intake of individual foods, nutrient intake from those foods, and the total nutrient intake by Americans. The nutrient intake data from NHANES, along with the dietary data from the surveys conducted by the Human Nutrition Information Service, U.S. Department of Agriculture (USDA), serve as the principal sources of national intake data for individuals (3–5,7). As such, nutritional intake data from NHANES are a key element in the National Nutrition Monitoring System (1,3,4). Selection of a dietary survey methodology or methodologies for use in an NHANES is, therefore, of fundamental importance.

Following the workshop, as stated in the introduction, detailed consideration was given to the workshop papers, presentations, and consensus statements. During several months following the workshop, additional planning sessions were held at NCHS that included survey planning staff and invited guests from NCHS, USDA, and other Public Health Service agencies. These included a weekly seminar series over 4 months that examined in detail individual topics related to the selection of a dietary survey methodology for use in the third NHANES (NHANES III). Additionally, the 1984 recommendations of the National Academy of Science's Coordinating Committee on Evaluation of Food Consumption Surveys were seriously considered, particularly the recommendations for linkage with USDA's Nationwide Food Consumption Survey (7). The recommendations of other expert panels were also reviewed and considered (8–10).

The results of this process were a series of decisions about which dietary assessment methods would be included in the survey. Food frequency instruments would be used to collect information on water and alcohol consumption, the intake of selected foods and food groups, food sources of calcium and vitamins A and C, and historical milk intake; the primary dietary survey instrument in NHANES III, as in past NHANES, would be the 24-hour recall. The decision to use the 24-hour recall as the primary dietary survey instrument was not a foregone conclusion, nor was it an easy decision.

In this chapter we will briefly explain the logic behind the decisions. Toward that goal we will discuss: Major uses of national nutrition data, specific information needs and constraints for NHANES III, types of dietary survey methods available, and a comparison of the availability of specific information from the 24-hour recall and the food frequency. Finally, we will describe in more detail the dietary survey methods being used in NHANES III.

Uses of nutrition data

In general, there are four major uses of national nutrition data: assessment and monitoring, regulatory uses, epidemiologic research, and commercial uses. The topic areas of the workshop were chosen with these uses in mind. Cancer, cardiovascular disease, energy balance (obesity), and osteoporosis were chosen because of their public

Table 1. Description of National Center for Health Statistics examination surveys

Survey	*Dates*	*Ages surveyed*	*Primary dietary survey instrument*
NHES I	1960–62	18–79 years	(1)
NHES II	1963–65	6–11 years	(1)
NHES III	1966–70	12–17 years	(1)
NHANES I	1971–75	1–74 years	24-hour recall
NHANES II	1976–80	6 months–74 years	24-hour recall
HHANES	1982–84	6 months–74 years	24-hour recall
NHANES I Followup	1982–present	25–74 years	Food frequency
NHANES III	1988–94	2 months and over	24-hour recall

[1]No dietary intake information was collected.

NOTES: NHES is National Health Examination Survey. NHANES is National Health and Nutrition Examination Survey. HHANES is Hispanic Health and Nutrition Examination Survey.

health importance. Moreover, they are especially related to nutrition. Generally, they are also highly prevalent diseases or conditions that can be detected both cross-sectionally and longitudinally. Although the topics appear to be weighted toward specific diseases and epidemiologic research, it must be remembered that none of the major uses are mutually exclusive. Epidemiologic data are an important component of the government's health and regulatory policy. Accordingly, for diet, the focus is on the ability to estimate absolute levels of intake rather than relative risk. Absolute levels of intake are essential for setting health guidelines or policy and for evaluating the magnitude (prevalence) of a purported health problem or outcome.

In the category "assessment and monitoring" the data collected would be used to produce national reference data ((1,2) and chapter 6). Questions to be addressed include: What is the mean intake for the population at large and for various selected subgroups (for example, sex, ethnic groups, and age)? What is the range of intakes? The dietary data will also be used to estimate the prevalence of intakes above or below some fixed cutpoint. In the past, for example, the Recommended Dietary Allowances (11) and the recommendations of the National Cholesterol Education Program (12) have been used to set these cutpoints. The results of such analyses can then be used to identify the groups within the country who are at greatest risk of overnutrition and undernutrition. With repeated NHANES, it will be possible to monitor trends in dietary intake over time. The trend data are especially useful both in evaluating the need for national nutrition programs and in monitoring the progress in meeting national nutrition and health objectives.

The principal reason for conducting NHANES is to supply U.S. Government agencies with information that will be useful in formulating and evaluating health policy and that will be the basis for regulatory actions. Government regulatory agencies ((13) and chapter 6) use survey data to determine and document the need for regulatory action, design a regulatory action, and evaluate the effectiveness of a particular regulatory action.

The impact of diet on the risk of certain conditions, disease, or even death can be evaluated using NHANES data. Such epidemiologic research can be cross-sectional in nature, as is typified by the role of NHANES data in exploring the association between dietary calcium intake and blood pressure (14–17). It can consist of an examination of aggregate trends over time (18), and with the completion of the NHANES I Epidemiologic Followup Survey (19,20), NHANES data have been used to examine the longitudinal relationships of diet to the subsequent risk of disease or death (21).

For the first time in the history of NHANES, NHANES III was designed to be both a cross-sectional and a longitudinal survey (1,5). At a minimum, NHANES III data will be linked to the National Death Index (22,23) in order to eventually ascertain the cause of death for all decedents. The objective will then be to relate baseline dietary intake and nutritional status to risk of death. The survey was also designed to have the potential for reinterviews and reexaminations. Thus, NHANES can no longer be thought of as only a cross-sectional survey.

There is a great deal of overlap in the uses of NHANES data. Assessment and monitoring uses are not distinct from regulatory or epidemiologic uses. A little recognized use of NHANES data that cuts across all three areas is the use of national prevalence data to make estimates of attributable risk (24).

Commercial interest in national nutrition data is, in theory, similar to the Government's nutrition monitoring interests. For example, What are the nutritional contributions of specific food items? When, where, what, how, how much, and with whom are meals consumed? Are there specific classes of foods which predominate in the food consumption patterns of specific subpopulations? Such information can then be used, for example, to develop advertising campaigns or new food products.

Information needs and constraints

In designing a survey like NHANES III, specific information must be collected if the uses outlined above are to be realized; very real logistical constraints limit what can be collected.

Information is needed about:

- The consumption of separate food items served individually, as components of mixed dishes, and together with other foods as meals.
- Food preparation methods, including type and amount of fat and salt added in cooking or at the table.
- Type of food—for example, fresh, frozen, canned, or dried.
- Brand names whenever possible (such as for infant formula, snacks, and prepared foods and meals).
- Sources of food—home, work, commercial, and so forth.
- Time of day.
- Water consumption.
- Usage of vitamin-mineral supplements, including brand name, dose, and frequency of use.
- Alcohol consumption.
- "Usual" or average intake of foods by individuals.
- Meaningful absolute values for intake, requiring the collection of exact portion size information.

Our objective was to select the dietary survey methodologies that would supply as much of this specific information as possible while also fitting within operational and methodological constraints. Perhaps the greatest constraints were the operational ones. In all cycles of NHANES, the participant is interviewed in the home approximately 2 to 3 weeks prior to his or her appointment at the mobile examination center (MEC) to undergo the examination procedures. The MEC appointment has historically included the dietary interview. Because of the time lag between contacts with the participant and the considerations mentioned above, the 24-hour recall has been used as the primary dietary survey methodology in NHANES.

Once in the MEC, the dietary interview can take no longer than 30 minutes before it impinges on the time available for the other tests and procedures included in the examination. To reduce data-processing errors, the information collected needed to be amenable to computerized collection. Indeed, NHANES III was planned to be the first survey in which an onboard computer system would be used to facilitate household sampling as well as to record the data collected during the examinations.

Consideration was also given to the ability of the different dietary methods to produce data comparable with that collected in past surveys, the method's appropriateness with respect to the research questions, its ability to serve as the basis for longitudinal followup studies, and the feasibility of administering questionnaires during the household interview to provide information for nonresponse bias analyses.

Types of dietary survey methods

There are two basic classes or types of dietary survey methods (table 2) available to choose from: daily food consumption methods and recalled "usual" or average food consumption methods (9,10).

The food record and the 24-hour recall are the most commonly used daily food consumption methods. Food records are considered by many to be the standard survey method for estimating dietary intake (25). Food records, however, have never been used in NHANES for several reasons. First, they require the training and active participation of the study participant regardless of whether food is weighed or merely recorded. Because of the level of effort required by the participant, there is concern about response rates and the comparability of recording skills among participants. The level of effort required can also lead participants to change their dietary patterns during the recording period. Recording abilities and completeness of the record can then affect dietary intake results. Moreover, the amount of time required to check the food records and to review them with the participant as well as the amount of effort required to code and process 30,000 food records, would be prohibitive in a survey the size of NHANES III.

Recalled "usual" or average food consumption methods comprise the other possible class of methods available (table 2). The dietary history as originally proposed by Bertha Burke (26,27) requires an extensive interview, which may take 1–2 hours to complete (25,27–30). In NHANES III about 20–30 minutes were allotted for the dietary interview in the MEC. Of the "usual" food consumption methods, the food frequency was the most likely choice based on time considerations in the MEC. There are two basic types of food frequencies, depending upon whether portion size information is collected during the interview (25,31). Because the NHANES III goal is to estimate both food and nutrient intakes, portion size information would also need to be specified on the recording form of the food frequency. Accordingly, the semiquantitative food frequency was the most likely choice of method from the "usual" class of methods.

Contrasting the availability of the specific information needed from the 24-hour recall and the semiquantitative food frequency, it was our judgment and the consensus recommendations of the expert consultants who participated in the workshop that a 24-hour recall be retained as the primary instrument for measuring dietary intake in NHANES III (chapter 7).

The availability of the necessary specific information from both the 24-hour recall and the food frequency is shown in table 3. This table is an adaption of a similar table given in the Life Sciences Research Office Report (9, p. 4).

Food frequency instruments, regardless of whether portion size information is collected, are usually precoded forms, consisting of a few selected "core" foods that are grouped into food groups of similar nutrient composition for a few selected key nutrients (25,32,33). Precoded forms are not new to the nutrition field. The concept was originally introduced by Donelson and Leichsenring in

Table 2. Types of dietary survey methods

Daily food consumption methods
Food record:
Weighed food record
Estimated portion sizes
24-hour recall
Recalled "usual" or average food consumption methods
Diet history:
Food frequency
24-hour recall
"Usual" meal patterns
Food frequency
Portion sizes unspecified or not noted
Portion sizes specified (also called "semiquantitative food frequency")

Table 3. Availability of specific information, by dietary survey methodology

Specific information	*24-hour recall*	*Semiquantitative food frequency*
Individual foods	Yes	No meals, few mixtures
Food preparation methods	Detailed	General
Type of food	Yes	Sometimes
Brand names	Some	Rare
Source	Yes	No
Time of day	Yes	No
Water consumption[1]	Yes	Yes
Supplements[1]	Yes	Yes
Alcohol	Less accurate	More accurate
"Usual" intake	Possible with replicates	Less accurate
Absolute intake	More accurate	Less accurate

[1]Asked as separate questionnaire items and not as part of 24-hour recall or food frequency in the third National Health and Nutrition Examination Survey.

SOURCE: Adapted from Life Sciences Research Office Guidelines for use of dietary intake data. Bethesda, Maryland: Federation of American Societies for Experimental Biology. 1986.

1942 as an effort to develop short-form methods for recording dietary records (34). Since that time their concept of the short form has been used to develop precoded food record forms (32,33,35,36) and the semiquantitative food frequency forms (37,38). Because the list of foods contains only a limited number of "core" foods and because foods are grouped together, food frequency instruments do not, as a rule, yield information comparable with that from a 24-hour recall.

Food frequencies do not generally yield intakes of individual foods, as mixtures or food eaten together as meals. They also do not supply information about food preparation methods, type of food, brand names, source of food, or time of day when the food was consumed.

Although both instruments have equivalent potential to estimate intake from water and vitamin-mineral supplements, food frequencies are likely to provide a better representation of alcohol consumption (chapter 3). Randomly collected multiple 24-hour recalls will provide, in our opinion, a better estimate of "usual" intake than food frequencies (chapter 1 and (9)).

Because food frequency forms often contain only a limited list of "core" foods, an open-ended section for recording foods not listed is desirable. If there is a great deal of cultural diversity, it will be impossible to develop a food list that includes all of the essential "core" foods eaten by different ethnic groups but at the same time is kept reasonably short. If there are too few foods listed, a great deal of effort will be needed to code other foods to an existing food group on the form or to create new food groups not listed on the form.

The food lists on a semiquantitative food frequency form are generally composed of single foods; as a result, an individual filling out a precoded form must separate most if not all salads and mixed dishes into their component foods. People must convert a recipe amount to a fraction of the standard portion size listed on the form, multiply that by the frequency with which that amount of a particular food is usually consumed, and then add to it the amounts of that food from other sources in order to estimate the usual consumption of the particular food. Having estimated the frequency of consumption of a single food, the results from all of the foods in the particular food group must then be added together to determine the total "usual" consumption of the food group. This is an extremely cumbersome and difficult process for most people to perform.

Once a food is recorded on a food frequency form, all the information about food preparation, foods consumed together as meals or mixed dishes, type of food, brand name, source of food, time of meals and snacks, ingredients, and portion size is lost. This point cannot be overemphasized or overstated. If there is ever a need to calculate the intake for a new dietary factor that was not a "key nutrient" in the development of the food groups on the food frequency form, it will be impossible to recombine foods into more appropriate food groups. Accordingly, food frequencies, and precoded records of any form, by definition must be used to meet only limited objectives. Many of the uses to which NHANES data have been put would have been impossible if a precoded semiquantitative food frequency—or food record, for that matter—based on a list of "core" foods had been used (see chapter 6 and (13)).

Methods selected for NHANES III

For the reasons outlined above, it is also our opinion that the 24-hour recall provides a more accurate estimate of the previous day's intake than a semiquantitative food frequency provides of the "usual" intake over the period of several months or a year (9,39). Accordingly, we accepted for NHANES III and have recommended to other countries planning national nutrition surveys the following:

- Use a daily food consumption method as the primary method for achieving "major use" needs.
- Collect at least 2 days of dietary intake information on at least a random sample of participants.
- Use a food frequency instrument to estimate water and alcohol intake, the frequency of consumption of selected foods and food groups of special interest, and information on historical intake of particular foods.

In NHANES III, a 24-hour recall interview will be conducted with every participant who comes to the MEC to undergo the examination procedures. Examinations will not be limited to a Tuesday–Saturday schedule, as in past Health and Nutrition Examination Surveys; therefore, 24-hour recall data will be collected on all days of the week. This will allow for the comparison of weekend and weekday dietary data and provide a more representative distribution of dietary intake data for all days of the week.

We plan to collect two additional 24-hour recalls by telephone for all examined sample persons 50 years of age and over. The National Institute on Aging of the National Institutes of Health has supported the followup telephone 24-hour recalls. Unfortunately, funding does not allow us to obtain multiple 24-hour recalls for all NHANES III participants, but a subsample of the examinees return for a replicate MEC examination and receive a second 24-hour recall.

A targeted food frequency is collected in NHANES III as part of the household interview. In addition, other questions on water and alcohol intake are asked in the MEC. The targeted food frequency will focus on food sources of calcium and vitamins A and C. The food frequency instrument does not ask the respondent to estimate "usual" portion size; accordingly, the food frequency data will not be used to produce nutrient intake estimates. Because the food frequency information is collected in the household, it will serve as a source of dietary information for all interviewed persons and will be used in followup analyses. The information will also be used as a further means to investigate response bias by comparing intakes of those participants who are interviewed only with intakes of those who are interviewed and examined. Finally, we are collecting historic frequency information on the intake of milk as a part of the osteoporosis component of NHANES III.

In summary, the key feature of any national dietary survey is flexibility in order to serve the multiplicity of data needs both within and outside the government. This point cannot be overstated. A current example of the need for flexibility is the introduction of new ingredients such as fat substitutes into the food supply. NHANES and the National Nutrition Monitoring System are designed first and foremost to meet the policy decisionmaking needs of the Federal Government. If NHANES dietary data are to serve a central role in nutrition assessment and monitoring, developing food policy and regulations related to food safety and fortification, and epidemiologic research for the Federal Government, it is essential to maintain as much flexibility as possible. In general, we believe that daily food consumption methods, in particular the 24-hour recall, provide this flexibility in NHANES III.

References

1. Woteki CE, Briefel RR, Kuczmarski R. Contributions of the National Center for Health Statistics. Am J Clin Nutr 47:320–8. 1988.
2. Murphy RS, Michael GA. Methodologic considerations of the National Health and Nutrition Examination Survey. Am J Clin Nutr 35:1255–8. 1982.
3. Woteki CE. Methods for surveying food habits: How do we know what Americans are eating? Clin Nutr 5:9–16. 1986.
4. Woteki CE. Dietary survey data: Sources and limits to interpretation. Nutr Rev suppl 204–13. 1986.
5. Murphy RS, Sempos C. Design considerations and dimensions of diet in the third National Health and Nutrition Examination Survey (NHANES III). In: Kohlmeier L, Helsing E, eds. Epidemiology, nutrition, and health. London: Smith-Gordon. 1989.
6. Yetley E, Johnson C. Nutritional applications of the Health and Nutrition Examination Surveys (HANES). Annu Rev Nutr 7:441–63. 1987.
7. National Academy of Sciences, National Research Council, Coordinating Committee on Evaluation of Food Consumption Surveys. National survey data on food consumption: Uses and recommendations. Washington: National Academy Press. 1984.
8. Life Sciences Research Office. Suggested measures of nutritional status and health conditions for the third National Health and Nutrition Examination Survey. Bethesda, Maryland: Federation of American Societies for Experimental Biology. 1985.
9. Life Sciences Research Office. Guidelines for use of dietary intake data. Bethesda, Maryland: Federation of American Societies for Experimental Biology. 1986.
10. Anderson SA. Guidelines for use of dietary intake data. J Am Diet Assoc 88:1258–60. 1988.
11. National Academy of Sciences, Subcommittee on the Tenth Edition of the RDAs. Recommended Dietary Allowances, tenth rev. ed. Washington: National Academy Press. 1989.
12. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Arch Intern Med 148:36–69. 1988.
13. Yetley EA, Hanson EA. Data sources and methods for estimating consumption of food components. J Toxicol Clin Toxicol 21(1 and 2):181–200. 1983–84.
14. Sempos C, Cooper R, Kovar MG, et al. Dietary calcium and blood pressure in National Health and Nutrition Examination Surveys I and II. Hypertension 8:1067–74. 1986.
15. Harlan WR, Hull A, Schmouder RL, et al. Blood pressure and nutrition in adults: The National Health and Nutrition Examination Survey. Am J Epidemiol 120:17–28. 1984.
16. McCarron DA, Morris CD, Henry HJ, Stanton JL. Blood pressure and nutrient intake in the United States. Science 224:1392–8. 1984.
17. Feinleib M, Lenfant C, Miller SA. Hypertension and calcium. Science 226:384–6. 1984.
18. Fulwood R, Rifkind B, Havlik R, et al. Trends in serum cholesterol levels among U.S. adults aged 20 to 74 years: Data from the National Health and Nutrition Examination Survey, 1960 to 1980. JAMA 257:937–42. 1987.
19. Cornoni-Huntley J, Barbano HE, Brody JA, et al. National Health and Nutrition Examination Survey I—Epidemiologic Follow-up Survey. Public Health Rep 98:245–51. 1983.
20. Cohen BB, Barbano HE, Cox CS, et al. Plan and operation of the NHANES I Epidemiologic Followup Study, 1982–84. National Center for Health Statistics. Vital Health Stat 1(22). 1987.
21. Jones DY, Schatzkin A, Green SB, et al. Dietary fat and breast cancer in the National Health and Nutrition Examination Survey I Follow-up Study. J Natl Cancer Inst 79:465–71. 1987.
22. Edlavitch SA, Baxter J. Comparability of mortality follow-up before and after the National Death Index. Am J Epidemiol 127:1164–78. 1988.
23. Patterson BH, Bilgrad R. Uses of the National Death Index in cancer studies. J Natl Cancer Inst 77:877–81. 1986.
24. Kahn HA, Sempos CT. Statistical methods in epidemiology. New York: Oxford University Press. 1989.
25. Becker BG, Indik BP, Beeuwkes AM. Dietary intake methodologies—a review. Ann Arbor: University of Michigan Research Institute. 1960.
26. Burke BS, Stuart HC. A method of diet analysis: Application in research and pediatric practice. J Pediatr 12:493–503. 1938.
27. Burke BS. The dietary history as a tool in research. J Am Diet Assoc 23:1041–6. 1947.
28. Truswell AS. How do we know what Australians eat? Med J Aust 149:2–4. 1988.
29. Marr JW. Individual dietary surveys: Purposes and methods. World Rev Nutr Diet 13:105–64. 1971.
30. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 115:492–505. 1982.
31. Sampson L. Food frequency questionnaires as a research instrument. Clin Nutr 4:171–8. 1985.
32. Johnson NE, Sempos CT, Elmer PE, et al. Development of a dietary intake monitoring system for nursing homes. J Am Diet Assoc 81:35–40. 1982.
33. Sempos CT, Johnson NE, Smith EL, et al. A two-year dietary survey of women. J Am Diet Assoc 84:1008–13. 1984.
34. Donelson EG, Leichsenring JM. A short method for dietary analysis. J Am Diet Assoc 18:429–34. 1942.

35. Johnson NE, Nitzke S, VandeBerg DL. A reporting system for nutrient adequacy. Home Econ Res J 2:210–21. 1974.
36. Hankin JH, Huenemann R. A short dietary method for epidemiologic studies. J Am Diet Assoc 50:487–92. 1967.
37. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65. 1985.
38. Block G, Hartman AM, Dresser CM, et al. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 124:453–69. 1986.
39. Liu K. Consideration of and compensation for intra-individual variability in nutrient intakes. In: Kohlmeier L, Helsing E, eds. Epidemiology, nutrition and health. London: Smith-Gordon. 1989.

Appendix I
NHANES III Dietary Survey Methodology Workshop program

Sunday, March 16, 1986

5:00 p.m. Registration
6:30 p.m. Dinner
8:00 p.m. Welcome—Robert Murphy, M.S.P.H.
8:15 p.m. Introduction to Planning for NHANES III—Kurt Maurer, Ph.D.
8:30 p.m. Nutrition Statistics and the Mission of NHANES III—Catherine E. Woteki, Ph.D., R.D.
Purpose of Workshop—Christopher T. Sempos, Ph.D.
Workshop Schedule and Procedures—Lenore Kohlmeier, Ph.D.
8:45 p.m. Break for the evening

Monday, March 17, 1986

8:00 a.m. Breakfast

Topic Number 1: Statistical issues related to survey design

8:45 a.m. Presenter: Kiang Liu, Ph.D. (Northwestern University)
9:15 a.m. First Discussant: Harold A. Kahn, M.A. (Consultant in epidemiology)
9:30 a.m. Second Discussant: George H. Beaton, Ph.D. (University of Toronto)
9:45 a.m. Open discussion
10:15 a.m. Break

Topic Number 2: Cancer

10:45 a.m. Presenter: Gladys Block, Ph.D. (National Cancer Institute)
11:15 a.m. First Discussant: Walter C. Willett, M.D., Dr. P.H. (Harvard University)
11:30 a.m. Second Discussant: Gail McKeown-Eyssen, Ph.D. (Ludwig Institute for Cancer Research)
11:45 a.m. Open discussion
12:15 p.m. Lunch

Topic Number 3: Cardiovascular disease

1:30 p.m. Presenter: Patricia J. Elmer, Ph.D., R.D. (University of Minnesota)
2:00 p.m. First Discussant: Arlene Caggiula, Ph.D., R.D. (University of Pittsburgh)
2:15 p.m. Second Discussant: Nancy Ernst, M.S., R.D. (National Heart, Lung, and Blood Institute)
2:30 p.m. Open discussion
3:00 p.m. Break

Topic Number 4: Energy balance

3:30 p.m. Presenter: Dorothy Blair, Ph.D. (Pennsylvania State University)
4:00 p.m. First Discussant: Johanna Dwyer, D.Sc., R.D. (New England Medical Center)
4:15 p.m. Second Discussant: Theodore Van Itallie, M.D. (Columbia University)
4:30 p.m. Open discussion
5:00 p.m. Adjournment
6:00 p.m. Social hour
7:00 p.m. Dinner

Tuesday, March 18, 1986

8:00 a.m. Breakfast

Topic Number 5: Osteoporosis

8:45 a.m. Presenter: Nancy E. Johnson, Ph.D. (University of Hawaii)
9:15 a.m. First Discussant: Mary Fran Sowers, Ph.D. (Cornell University)
9:30 a.m. Open discussion
10:00 a.m. Break

Topic Number 6: Nutrition monitoring

10:30 a.m. Presenter: Elizabeth A. Yetley, Ph.D., R.D. (Food and Drug Administration)
11:00 a.m. First Discussant: Eleanor M. Pao, Ph.D. (U.S. Department of Agriculture)
11:15 a.m. Second Discussant: Susan B. Foerster, M.P.H., R.D. (Association of State and Territorial Public Health Nutrition Directors)
11:30 a.m. Open discussion
12:15 p.m. Lunch

Dietary survey methods—An overview

1:30 p.m. Presenter: Lenore Kohlmeier, Ph.D. (Federal Health Office, Federal Republic of Germany)
2:15 p.m. Open discussion
3:00 p.m. Thank you—Robert S. Murphy, M.S.P.H. (National Center for Health Statistics)
3:10 p.m. Adjournment

Appendix II
NHANES III Dietary Survey Methodology Workshop Planning Committee

Christopher T. Sempos, Ph.D.
Cochairperson
Nutritional Statistics Branch
National Center for Health Statistics

Ronette R. Briefel, Dr.P.H., R.D.
Cochairperson
Survey Planning and Development Branch
National Center for Health Statistics

Lenore Kohlmeier, Ph.D.
Moderator
Institute for Social Medicine and Epidemiology
Federal Health Office
Federal Republic of Germany

Clifford Johnson, M.S.P.H.
Chief, Nutritional Statistics Branch
National Center for Health Statistics

Kurt R. Maurer, Ph.D.
Chief, Survey Planning and Development Branch
National Center for Health Statistics

Robert S. Murphy, M.S.P.H.
Director, Division of Health Examination Statistics
National Center for Health Statistics

Catherine E. Woteki, Ph.D., R.D.
Deputy Director, Division of Health Examination Statistics
National Center for Health Statistics

Appendix III Directory of workshop participants[1]

Presenters and discussants

Lenore Kohlmeier, Ph.D.
Department of Epidemiology of Health Risk
Institute for Social Medicine and Epidemiology
Federal Health Office
General Pape Strasse 64–66
D-1000, Berlin 42
Federal Republic of Germany

George H. Beaton, Ph.D.
Department of Nutritional Sciences
Faculty of Medicine
University of Toronto
Toronto, Ontario
Canada M5S 1A8

Dorothy Blair, Ph.D.
Nutrition Program
College of Human Development
Pennsylvania State University
Unviersity Park, PA 16802

Gladys Block, Ph.D.
National Cancer Institute, DCPC
EPN, Room 313
9000 Rockville Pike
Bethesda, MD 20892–4200

Arlene Caggiula, Ph.D., R.D.
Department of Epidemiology
Graduate School of Public Health
University of Pittsburgh
Pittsburgh, PA 15261

Johanna Dwyer, D.Sc., R.D.
Department of Medicine
Tufts University School of Medicine and New England Medical Center Hospitals
Frances Stern Nutrition Center
171 Harrison Avenue
Boston, MA 02111

Patricia J. Elmer, Ph.D., R.D.
Division of Epidemiology
University of Minnesota
Stadium Gate 27
611 Beacon Street, SE
Minneapolis, MN 55455

Nancy Ernst, M.S., R.D.
Division of Epidemiology and Clinical Applications-OD
National Heart, Lung, and Blood Institute
National Institutes of Health
Federal Building, Room 2A04
7550 Wisconsin Avenue
Bethesda, MD 20892–4200

Gail McKeown-Eyssen, Ph.D.
Ludwig Institute for Cancer Research
Toronto Branch
Toronto, Ontario
Canada M4Y 1M4

Susan B. Foerster, M.P.H., R.D.
California Department of Health Services
President, Association of State and Territorial Public Health Nutrition Directors
714 P Street
Sacramento, CA 95814

Nancy E. Johnson, Ph.D.
Department of Food Science and Human Nutrition
University of Hawaii
Henke Hall, Room 224
1800 East-West Road
Honolulu, HI 96822

Harold A. Kahn, M.A.
Consultant in Epidemiology
3405 Pendleton Drive
Silver Spring, MD 20902

Kiang Liu, Ph.D.
Department of Community Health and Preventive Medicine
Northwestern University
303 East Chicago Avenue
Chicago, IL 60611

[1]Represents affiliation at the time of the workshop.

Eleanor M. Pao, Ph.D.
Food Consumption Research Branch
Human Nutrition Information Service
U.S. Department of Agriculture
6505 Belcrest Road
Hyattsville, MD 20782

Mary Fran Sowers, Ph.D.
Division of Nutritional Sciences
Cornell University
302 MVR Hall
Ithaca, NY 14853

Theodore B. Van Itallie, M.D.
Department of Medicine
Columbia University College of Physicians and Surgeons
St. Luke's Roosevelt Hospital Center
Amsterdam Avenue and 114th Street
New York, NY 10025

Walter C. Willett, M.D., Dr.P.H.
Department of Epidemiology
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115

Elizabeth A. Yetley, Ph.D., R.D.
Center for Food Safety and Applied Nutrition
Food and Drug Administration
HFF-265
200 C Street, SW
Washington, DC 20204

Attendees

Melody Bacha, M.P.H., R.D.
Analysis and Evaluation
Food and Nutrition Service
U.S. Department of Agriculture
Room 1024
3101 Park Center Drive
Alexandria, VA 22302

Marilyn Buzzard, Ph.D., R.D.
Nutrition Coordinating Center
University of Minnesota
Suite 526
2829 University Avenue, SE
Minneapolis, MN 55414

Eva Callmer, Ph.D.
Department of Medical Nutrition
Karolinska Institute
Huddinge University Hospital F69
S-14186 Huddinge
Sweden

Carolyn Clifford, Ph.D.
National Cancer Institute
National Institutes of Health
Blair Building, Room 623
7550 Wisconsin Avenue
Bethesda, MD 20892–4200

Marcie Cynamon, M.S.
Division of Health Interview Statistics
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, MD 20782

Alta Engstrom, Ph.D.
Director, Nutrition Marketing Service
General Mills
Box 1113
Minneapolis, MN 55440

Donald Everett, M.S.
National Institute on Aging
National Institutes of Health
612 Federal Building
7550 Wisconsin Avenue
Bethesda, MD 20892

Marilyn Farrand, M.S., R.D.
National Heart, Lung, and Blood Institute
National Institutes of Health
6A08 Federal Building
7550 Wisconsin Avenue
Bethesda, MD 20892

Richard Havlik, M.D.
Office of Program Planning, Evaluation, and Coordination
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, MD 20782

John Heybach, Ph.D.
Director, Scientific Affairs
NutraSweet Group
4711 Golf Road
Skokie, IL 60076

Joan Cornoni-Huntley, Ph.D.
National Institute on Aging
National Institutes of Health
612 Federal Building
7550 Wisconsin Avenue
Bethesda, MD 20892

Mary Grace Kovar, Dr.P.H.
Office of Interview and Examination Statistics Program
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, MD 20782

Virginia Laukaran, Dr.P.H.
Food and Nutrition Board
National Research Council
2101 Constitution Avenue (NAS-340)
Washington, DC 20018

Christine Lewis, Ph.D., R.D.
Clinical Nutrition Branch
Food and Drug Administration, HFF-265
200 C Street, SW
Washington, DC 20204

Luize Light, Ed.D.
National Cancer Institute
National Institutes of Health
Blair Building, Room 420
7550 Wisconsin Avenue
Bethesda, MD 20892–4200

James T. Massey, Ph.D.
Office of Research and Methodology
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, MD 20782

Linda Meyers, Ph.D.
Office of Disease Prevention and Health Promotion, OASH, DHHS
Switzer Building, Room 2132
330 C Street, SW
Washington, DC 20201

Susan Pilch, Ph.D.
Federation of American Societies for Experimental Biology
Life Sciences Research Office
9650 Rockville Pike
Bethesda, MD 20814

William Rand, Ph.D.
INFOODS Network, Room 20A-226
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

Robert Rizek, Ph.D.
Nutrition Monitoring Division
Human Nutrition Information Service
U.S. Department of Agriculture
Federal Building, Room 368
6505 Belcrest Road
Hyattsville, MD 20782

Harold Roth, M.D.
Division of Digestive Diseases and Nutrition Epidemiology and Data Systems Program
National Institute of Arthritis, Diabetes, Digestive and Kidney Diseases
National Institutes of Health
Federal Building, Room 616
Bethesda, MD 20892

Kathleen Turczyn, M.S.
Division of Epidemiology and Health Promotion
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, MD 20782

John E. Vanderveen, Ph.D.
Division of Nutrition, HFF-260
Food and Drug Administration
200 C Street, SW
Washington, DC 20204

Regina Ziegler, Ph.D.
National Cancer Institute
National Institutes of Health
Environmental Epidemiology Branch
Landow Building, Room 3C07
7910 Woodmont Avenue
Bethesda, MD 20892

Division of Health Examination Statistics:

Ronette R. Briefel, Dr.P.H., R.D.
Stacey FitzSimmons, Ph.D.
Dale Hitchcock, B.S.
Clifford Johnson, M.S.P.H.
Robert Kuczmarski, Dr.P.H., R.D.
Andrea LaCroix, Ph.D.
Anne Looker, Ph.D., R.D.
Kurt R. Maurer, Ph.D.
Marie Mitchell, B.S.
Robert S. Murphy, M.S.P.H.
Christopher T. Sempos, Ph.D.
Deborah Winn, Ph.D.
Catherine E. Woteki, Ph.D., R.D.

National Center for Health Statistics
6525 Belcrest Road
Hyattsville, MD 20782

Appendix IV Outline of statistical issues related to NHANES III dietary survey methodology

I. Review of dietary survey methodology in the past
  A. Methods used in specific study designs
    1. Case-control
    2. Cross-sectional (prevalence)
    3. Cohort (incidence)
  B. Problems with dietary surveys in the past as related to statistical issues
    1. Assumptions made in collection of data and study design
      a. Case-control
      b. Cross-sectional
      c. Cohort
    2. Number of records collected
    3. Uses of data in analysis

II. Within- and between-person variability in food consumption
  A. Definitions
    1. Usual intake
    2. Within-person variability
    3. Between-person variability
  B. Effects of large ratios of within- to between-person variability
    1. Case-control
      a. Based on or referring to baseline data
      b. Dietary history for some specified period of life prior to NHANES III taken at baseline
    2. Cross-sectional
    3. Cohort
  C. Estimation of number of dietary records to be collected per individual
    1. How should number be determined?
    2. Should number vary by age, race, sex?
  D. Effects of number of dietary records collected per individual on later analyses
    1. Group comparisons
      a. Analysis of variance
      b. Correlation
      c. Regression
        i. Simple
        ii. Multiple
      d. Measures of risk
        i. Odds ratio
        ii. Relative risk
    2. Individual comparisons
      a. Correlation
      b. Regression
        i. Simple
        ii. Multiple
      c. Measures of risk
        i. Odds ratio
        ii. Relative risk

III. Recommendations
  A. Should multiple dietary records be collected?
    1. How many?
    2. When should they be collected?
  B. Limitations of recommendations on study design for future analyses

Appendix V
Outline of dietary methodology issues for presenters

I. Review of dietary methods related to (<u>topic area</u>) in the following groups
   A. Adults
   B. Adolescents
   C. Infants and children

II. Critique of different methods
   A. Ability of different dietary survey methods to estimate various dietary constituents of concern for topic area

      Cancer
      1. Total calories
      2. Fat
      3. Vitamin A
      4. Carotenoids
      5. Selenium
      6. Vitamin E
      7. Vitamin C
      8. Fiber

      Cardiovascular disease
      1 Total calories
      2. Fat
      3. Cholesterol
      4. Sodium
      5. Potassium
      6. Calcium
      7. Caffeine
      8. Alcohol

      Energy balance
      1. Total calories
      2. Fat
      3. Protein
      4. Carbohydrate
      5. Alcohol

      Osteoporosis
      1. Total calories
      2. Protein
      3. Calcium
      4. Phosphorus
      5. Fluoride
      6. Magnesium
      7. Water
      8. Vitamin D

      Nutrition monitoring
      1. Total calories
      2. Iron, sodium, and other minerals
      3. Vitamin A and other fat-soluble vitamins
      4. Folacin and other water-soluble vitamins
      5. Vitamin-mineral supplements
      6. Nonnutritional dietary components
   B. Applicability to various study designs
      1. Case-control
      2. Cross-sectional (prevalence)
      3. Cohort (incidence)

III. Recommendations
   A. Method(s) recommended
      1. Case-control
         a. Based on or referring to baseline data
         b. Dietary history for some specified period of life prior to NHANES III taken at baseline
      2. Cross-sectional
      3. Cohort
   B. Application to NHANES
      1. Time limit for dietary interview of 30 minutes
      2. Continuity and comparability with past and present national nutrition surveys
         a. NHANES
         b. Nationwide Food Consumption Survey (U.S. Department of Agriculture)
      3. Potential for automation
      4. Equipment needed
      5. Training for interviewers
   C. Validity of method(s) recommended
   D. Should multiple records be collected?
      1. How many?
      2. Using which dietary survey methods? (if recommendation is to use two or more dietary survey methods)
   E. When should multiple records be collected?
      1. Prior to medical examination
         a. Household interview
         b. Other
      2. Medical examination
      3. After medical examination
   F. Method of collection of multiple records
      1. Interview in person
      2. Telephone interview
      3. Mailed interview

Appendix VI Prospectus for workshop

The third National Health and Nutrition Examination Survey (NHANES III) will be fielded in 1988. At present, the National Center for Health Statistics is soliciting suggestions for survey content and topics to be considered for NHANES III. It is anticipated that a 3-hour examination consisting of questionnaires, biochemical measurements, and medical procedures will be administered to approximately 30,000 persons ages 2 months and older over a 6-year period. The examination phase will address certain target conditions. Components under consideration for NHANES III include overweight and obesity, cardiovascular disease, osteoporosis, cancer, and nutrition monitoring.

Objectives

The major purpose of the workshop is to review, evaluate, and make recommendations regarding existing and potential dietary methodologies with respect to target conditions to be studied in NHANES III. For the purposes of this conference, dietary intake methodologies will be evaluated with respect to proposed target conditions that include a nutrition component as well as to the overall objectives of NHANES III:

1. To monitor changes in health and nutritional status over time.
2. To assess the health and nutritional status of the U.S. population and specific subgroups.
3. To provide information on the interrelationship of health and nutrition variables within population subgroups.
4. To measure the prevalence of disease and conditions.
5. To measure met and unmet care needs related to target conditions.

Data needs

Previous Health and Nutrition Examination Surveys have utilized various methodologies (24-hour recall, food frequency, and questionnaires) to describe the dietary intake and nutritional habits of the U.S. population. The 24-hour recall yields quantitative estimates of recent dietary intake for groups, whereas multiple 24-hour recalls, food records, or a diet history provide quantitative intake for individuals. The diet history and the food frequency methods yield estimates of qualitative intake over time for both individuals and groups and, therefore, may be appropriate methods to use for investigating the relationship between past dietary intake and health. Food frequencies can be targeted to specific foods, food groups, or nutrients depending on research needs. Quantitative data can be estimated from food frequency data using portion size information, but a standardized food composition data base for use with frequency questionnaires has not been developed.

Current dietary methodologies used in clinical studies and surveys are appropriate for different research situations. The traditional dietary methods can be time consuming and costly in terms of respondent time, interviewer expertise required, and recording and editing of data; thus, they may not be practical for use in a large survey such as NHANES III. In recent years, epidemiologic research has been aimed at developing and testing other dietary methods, such as the semiquantitative food frequency, to provide dietary data for both individuals and groups.

The National Nutrition Monitoring System requires that national estimates of dietary intake continue to be provided. The Food and Drug Administration monitors the nutritional adequacy of the population's intake for food fortification policy. In addition to the continuing need to provide descriptive data (for example, the prevalence of overweight and obesity and the mean dietary intake of specific population groups), a growing need exists for information in the applied nutrition area to test specific research questions. The collection of accurate dietary intake data for individuals is desirable for comprehensive research analyses relating nutrition to health and specific target conditions.

In order to examine the interactions of nutrition-related variables with health conditions or to relate food consumption patterns and nutrient intake to physiological and physical indicators of health status, the dietary methodologies used must be appropriate for the target condition, applicable at the individual level, and practical in regard to the survey design. If NHANES III uses a longitudinal design, adequate baseline dietary and

NOTE: Provided as background information to workshop participants prior to the workshop.

nutritional data that relate to the followup of particular target conditions must be collected in order to test specific hypotheses regarding nutrition and health status. Therefore, the collection of reliable and valid data for both individuals and groups of individuals will most likely be required for NHANES III.

Issues to be considered when determining the dietary methodologies for NHANES III include: (1) the survey design; (2) the practicality of the dietary method in terms of use for a large survey, computer automation, the length of the interview, and cost; (3) the comparability to previous methods used in NHANES; (4) the 1983 recommendations of the Coordinating Committee on the Evaluation of Food Consumption Surveys with respect to linkage with the Nationwide Food Consumption Survey; (5) the appropriateness for different population groups; (6) interviewer training; and (7) the representativeness of an individual's usual intake.

The survey's data needs can be interpreted in terms of whether the analytical research requires precise quantitative data on nutrient intake or the relative classification or ranking of individuals with respect to nutrient intake. Consideration must also be given to whether the data needed relate to cross-sectional estimates of recent group intake, estimates of past or long-term individual intake, or both. Conferences on dietary methodology have been previously conducted, but none has dealt with the issues that relate specifically to the data needs, limitations, and objectives of NHANES.

Proposed approach

The NHANES III planning schedule includes the solicitation of project proposals and topic review prior to the dietary conference. Target conditions will be defined during late 1985. The conference will be held March 16–18, 1986. Each speaker will be asked to prepare a background paper related to one of the major target components or statistical issues. In addition, two expert consultants for each of the components will comment on the background papers and also participate in the conference. Thus, the conference will involve approximately 18 outside consultants, as well as NHANES III planning staff, and representatives from the U.S. Department of Agriculture, the Food and Drug Administration, and the National Institutes of Health, for a total of approximately 40 participants. Each speaker will be given an outline to be used in the preparation of the background papers explaining the specific issues to be addressed and requesting suggestions and recommendations for dietary methodologies to be used in NHANES III. After discussions held during the conference, a revised background paper will be prepared by each of the six speakers, in consultation with the discussants. Finally, each of the six working groups will prepare a consensus statement recommending dietary methodologies to be used in NHANES III for the group's specific topic area.

Appendix VII NHANES III survey design

The National Center for Health Statistics is planning the third National Health and Nutrition Examination Survey (NHANES III). NHANES I took place in 1971–75 and NHANES II, in 1976–80. As in previous NHANES programs, the primary purpose of NHANES III will be to measure and monitor the health and nutritional status of the U.S. population. In addition, projects of an analytic nature are likely to be included in NHANES III, and serious thought is also being given to the establishment of a cohort study using NHANES III as the baseline. This cohort study may include reexamination at intervals of every 6 years, mail followup, and monitoring of deaths through the National Death Index.

NHANES III, like previous NHANES programs, will consist of a household medical history interview and a medical examination conducted in a mobile examination center 1 to 4 weeks after the household interview. It is anticipated that the 3-hour medical examination will consist of questionnaires, biochemical measurements, and medical procedures. In the past the examination has been administered to approximately 20,000 persons ages 6 months–74 years over a 4-year period. In NHANES III, we plan to interview 40,000 persons and to examine 30,000 persons ages 2 months and older during a 6-year period to improve coverage of black and Hispanic persons and to permit time-trend analysis.

The dietary interview in the past has been conducted during the 3-hour medical examination. In previous NHANES it consisted of a 24-hour recall and food frequency interview covering the past 3 months. The entire dietary interview took approximately 30 minutes.

If we are to survey 40,000 persons over a 6-year period, it is essential that we plan very carefully the dietary survey methods to be used in NHANES III so that we can participate effectively in nutrition monitoring efforts; study the relationship of diet to specific target conditions; maintain comparability with past NHANES and the U.S. Department of Agriculture's Nationwide Food Consumption Survey; and automate as much of the dietary data collection as possible.

NOTE: Provided as background information to workshop participants prior to the workshop.

Appendix VIII Measuring dietary patterns in surveys

by Catherine E. Woteki, Ph.D., R.D., National Center for Health Statistics

Introduction

In nutrition surveys, five types of measurements are needed to completely characterize the nutritional status of a person or a population: Food and nutrient intakes, body measurements, hematological and biochemical tests, physical examination for the presence of clinical signs of deficiency or toxicity, and medical history. No one method alone is sufficient for assessing the nutritional status of an individual or a group of individuals.

In approaching the design of a survey in which nutrition is to be assessed, the selection of which nutritional assessment methods are to be used depends on the answers to two questions: What aspects of nutrition are of interest? How suitable are the available measures? Clear statements of the goals of the survey and of hypotheses to be tested are helpful in deciding which aspects of nutrition are of interest and which broad categories of measurements should be included. Consideration as to the practicality, reliability, and validity of available measurements will determine their suitability for the survey under design.

In a recent article, Ware et al. (1) discuss issues of practicality, reliability, and validity when selecting measures of health status. Their discussions are also pertinent to selection of nutritional status indicators. The authors caution that health status, like nutritional status, cannot be observed directly and that investigators can only make inferences about health or nutritional status from fallible indicators. Whether a given measurement is practical depends upon the total measurement resources available, respondent burden, and analytical resources. Total measurement resources and the priority assigned to nutritional assessment will to a large degree determine the amount of staff time and resources to be dedicated to nutritional assessment. The burden on the respondent in terms of time, inconvenience, and physical discomfort will affect refusal rates, rates of missing responses, and administration time in followup of nonrespondents. The degree of reliability needed depends on the purpose of the study, with more sensitive and specific measures required for clinical decisionmaking than for the comparison of groups of people. Validity of nutritional status indicators should be evaluated within the context of the particular health or nutritional components of interest to the study. The literature with respect to reliability and validity of nutritional status measurements is limited, offering opportunities for future research.

This appendix is a description of the data available from national nutrition surveys and a review of the strengths, weaknesses, and limits of interpretation. The nutrient content of the food supply and the 24-hour recall of food consumption are discussed in detail.

Data available from national surveys

The Federal Government collects a broad range of nutrition-related information under the general heading of the National Nutrition Monitoring System (NNMS). The Food and Agriculture Act of 1977 (Public Law 95–113) instructed the Secretary of Agriculture and the Secretary of Health, Education, and Welfare to submit to Congress a proposal for a comprehensive nutritional status monitoring system that would integrate the ongoing nutrition survey activities of both Departments. The Departments' original proposal was submitted to Congress in May 1978, and at the request of the Committee on Science and Technology, the General Accounting Office (GAO) reviewed it. In June 1978, GAO sent a letter to the Secretaries of the two Departments recommending the development of a comprehensive implementation plan. The Joint Implementation Plan for a Comprehensive National Nutrition Monitoring System was submitted to Congress in September 1981. Ostenso (2) has described in greater detail the history of nutrition monitoring.

The Implementation Plan classified the activities conducted under NNMS into five major categories:

- Health status measurements.
- Food consumption measurements.
- Food composition.
- Dietary knowledge and attitudes.
- Food supply determinations.

The activities in each of these categories and the agencies responsible for them are described below.

Health status measurements

The Department of Health and Human Services (DHHS), through the National Center for Health Statistics (NCHS), the Centers for Disease Control (CDC), and other agencies, collects a broad array of data on the health status of the American population. Two programs are specifically oriented to nutrition. These are the National Health and Nutrition Examination Surveys (NHANES), conducted by NCHS, and the Coordinated State Surveillance System (CSSS), conducted by CDC. Table 1 summarizes the specific nutrition-related indicators for which data are available from NHANES and CSSS.

NHANES, conducted periodically by NCHS, is the cornerstone of Federal efforts to monitor the overall nutritional status of the American people. NHANES consists of a series of surveys carried out on a representative sample of the civilian noninstitutionalized U.S. population, comprising more than 20,000 persons in each survey. The surveys, which include health histories, dietary interviews, physical examinations, and laboratory measurements, provide information on national health and nutritional status. Two national surveys have been completed—NHANES I (1971–75) and NHANES II (1976–80)—as well as a survey of the health and nutritional status of Hispanic Americans, conducted in 1982–84. Descriptions of the plan and operation of these surveys have been published by NCHS (3–6).

Through NHANES, physical and biochemical measurements are made that provide information about a number of nutrition-related conditions, including growth retardation, anemia, obesity, heart disease, hypertension, diabetes mellitus, vitamin and mineral deficiency or toxicity, and heavy metal and pesticide exposure. The types of measurements include physical examinations, anthropometry, hematological assessments, biochemical analyses of blood and urine, x rays, and functional assessments, as well as a health history.

The Nutrition Division, Center for Chronic Disease Prevention and Health Promotion, CDC, contributes to nutrition monitoring through a nutrition surveillance program. CSSS monitors the nutritional status of high-risk pediatric and pregnant populations through the collection of measurements readily available, such as height, weight, hemoglobin, and hematocrit. The system uses information from service delivery programs operated by selected State and metropolitan health jurisdictions to provide data on the prevalence of major nutritional problems in the targeted groups. The composition of the groups under surveillance is determined by their socioeconomic status, their proximity to a nonprivate outpatient clinic, and the fact that selected State health departments are cooperating with the CDC program in obtaining and utilizing nutrition-related data. The indicators of health status measured in CSSS are limited to a relatively few indexes related to nutritional problems identified in the NHANES activities. CSSS provides information about the prevalence of overweight, underweight, retarded linear growth, and anemia among high-risk children. In addition, pregnant women are kept under surveillance with attention to such indicators as anemia, abnormal weight changes, fetal survival, birth weight of the child, and whether breast or bottle feeding is used and for how long. The CDC publishes annual reports on its findings (7).

Table 1. Sources of data on indicators of nutritional status

Indicator	*National Health and Nutrition Examination Survey*		*Coordinated State Surveillance System*	
	I[1]	*II*[2]	*Pediatric*	*Pregnant*
Height	Yes	Yes	Yes	–
Weight	Yes	Yes	Yes	–
Skinfold thickness:				
Triceps	Yes	Yes	–	–
Subscapular	Yes	Yes	–	–
Head circumference	–	Yes	Yes	–
Hemoglobin	Yes	Yes	Yes	Yes
Hematocrit	Yes	Yes	Yes	Yes
Red blood cell count	–	Yes	–	–
White blood cell count	Yes	Yes	–	–
Mean corpuscular volume	---	Yes	---	---
Mean corpuscular hemoglobin concentration	---	Yes	---	---
Red blood cell protoporphyrin	---	Yes	---	---
Serum iron	Yes	Yes	---	---
Serum total iron-binding capacity	Yes	Yes	---	---
Transferrin saturation	Yes	3–74 years	---	---
Ferritin	---	([3])	---	---
Serum zinc	---	3–74 years	---	---
Serum copper	---	3–74 years	---	---
Serum vitamin C	---	3–74 years	---	---
Serum albumin	---	3–74 years	---	---
Serum vitamin A	---	3–11 years	---	---
Red blood cell folate	---	([3])	---	---
Serum folate	---	([3])	---	---
Serum cholesterol	Yes	20–74 years	---	---
Serum triglycerides	---	20–74 years	---	---
High-density lipoprotein	---	20–74 years	---	---
Serum vitamin B_{12}	---	([3])	---	---
Breastfeeding	---	([3])	---	Yes
Low birth weight	---	---	Yes	Yes

[1]Ages surveyed are 1–74 years.
[2]Unless otherwise specified, ages surveyed are 6 months–74 years.
[3]Performed on a subsample of persons ages 3–74 years.

Food consumption measurements

Four surveys provide information on the food people consume. They are the Nationwide Food Consumption Survey (NFCS) and the Continuing Survey of Food Intakes by Individuals (CSFII), both of which are conducted by the Human Nutrition Information Service (HNIS) of the Department of Agriculture (USDA), the NHANES (previously described), and the Food Usage Survey conducted by the Food and Drug Administration (FDA). Table 2 shows the nutrients for which estimates are made in NFCS, CSFII, NHANES, and the national food supply determination (discussed later).

NFCS has been conducted at roughly 10-year intervals since the mid-1930's. It is actually two surveys: A survey of household food use, which has been conducted six times, and a survey of individuals' food intakes, which has been conducted twice. The most recent NFCS was conducted in 1977–78 and consisted of a basic survey of 15,000 households and household members and five additional surveys—surveys conducted in Alaska, Hawaii, Puerto Rico and surveys in the continental United States of low-income and elderly households. In NFCS, the method of collecting household data is a 7-day recall of food used.

Table 2. Sources of data on nutrients and other food constituents

Nutrient or food constituent[1]	National food supply	Nationwide Food Consumption Survey			NHANES		Total Diet Study	Food composition
		Household	Individual	CSFII[2]	I	II		
Water	---	---	---	Yes	---	---	---	Yes
Energy (kcal)	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Protein:								
Total	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Amino acids	---	---	---	---	---	---	---	([3])
Carbohydrate:								
Total	Yes	Yes	Yes	Yes	---	Yes	---	Yes
Sugars	Yes	---	([4])	---	---	---	---	---
Lipids:								
Total fat	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Saturated fat	Yes	---	---	Yes	Yes	Yes	---	([3])
Oleic acid	Yes	---	---	---	Yes	Yes	---	([3])
Total monounsaturated	---	---	---	Yes	---	---	---	([3])
Linoleic acid	Yes	---	---	---	Yes	Yes	---	([3])
Total polyunsaturated	---	---	---	Yes	---	---	---	([3])
Cholesterol	Yes	---	([4])	Yes	Yes	Yes	---	([3])
Vitamins:								
A value, international units	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
A value, retinol equivalents	---	---	---	Yes	---	---	---	([3])
Carotene	---	---	---	Yes	---	---	---	---
E	---	---	---	Yes	---	---	---	([3])
Thiamine (B_1)	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Riboflavin (B_2)	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Niacin (preformed)	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Pantothenic acid	([4])	---	---	---	---	---	---	([3])
B_6	Yes	Yes	Yes	Yes	---	---	---	Yes
Folic acid	([4])	---	---	Yes	---	---	---	([3])
B_{12}	Yes	Yes	Yes	Yes	---	---	---	Yes
C	Yes	Yes	Yes	Yes	Yes	Yes	---	Yes
Minerals:								
Calcium	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Phosphorus	Yes	Yes	Yes	Yes	---	Yes	Yes	Yes
Magnesium	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Iron	---	---	---	---	---	---	Yes	---
Iodine	---	---	---	---	---	---	Yes	---
Sodium	Yes	---	Yes	Yes	---	Yes	Yes	Yes
Potassium	Yes	---	---	Yes	---	Yes	Yes	Yes
Copper	---	---	---	Yes	---	---	Yes	([3])
Zinc	Yes	---	---	Yes	---	---	Yes	([3])
Manganese	---	---	---	---	---	---	Yes	([3])
Selenium	---	---	---	---	---	---	Yes	---
Chromium	---	---	---	---	---	---	Yes	---
Fiber:								
Crude	Yes	---	---	---	---	---	---	Yes
Dietary	---	---	---	Yes	---	---	---	Yes
Alcoholic beverages	---	Yes	Yes	Yes	---	Yes	Yes	Yes

[1]Alcoholic beverages are included.
[2]Continuing Survey of Food Intakes by Individuals
[3]Nutrient data available at the completion of revision of U.S. Department of Agriculture's Handbook No. 8.
[4]SOURCE: Food Supply and Nationwide Food Consumption Survey data supplied by Dr. Susan Welsh, Human Nutrition Information Service, U.S. Department of Agriculture.

The individual household members recall 1 day's food intake and keep a diary for 2 additional days. Descriptions of the survey design and results are published (8,9) and also made available as computer tapes for researchers.

HNIS also conducts CSFII as a component of the National Nutrition Monitoring System (2). The first CSFII collected 1-day food and nutrient intake for women 19–50 years of age and their children ages 1–5 years in the 48 conterminous States during April and June 1985 (10). This sample, the "core monitoring group," was selected as the group shown to be at nutritional risk in previous surveys. The survey will be conducted annually and additional population groups may be added in the future.

Using commercial market research data bases (A.C. Nielsen Company), FDA conducts, on a biennial basis, a survey of a statistically representative sample of products representing major food classes from the total packaged food supply. The Food Usage Survey involves approximately 1,700 individual food brands representing about 44 percent of the packaged food supply in retail dollar terms, which in turn is generalizable to the total packaged food supply. The ingredient data are the basis for the FDA food ingredient data bank and are used for multiple special studies. Significant changes in aggregated public purchasing patterns and the food industry's reactions can be followed. For example, changes in public purchasing practices associated with avoiding specific components of foods (such as sugars and food additives) may be quickly identified. The same data base is used to measure changes in aggregate public purchasing practices of nutritionally modified foods, such as fortified foods, low-sodium and reduced-sodium foods, low-cholesterol and reduced-calorie foods. Through this means, public responses to nutrition information and education programs, new labeling approaches, media coverage, and other societal events can be measured. This type of information permits estimates of the impact of programs initiated by government and the private sector in the interest of improving nutritional health.

Food composition

Assessing the nutritional adequacy of diets reported in NHANES and NFCS would be impossible without information on the nutrient content of foods. Four different activities contribute to our knowledge of food composition: The Nutrient Data Bank, the Total Diet Study, Labeled Food Surveillance, and research.

The USDA Nutrient Data Bank (NDB) is the major mechanism for collecting, evaluating, storing, and collating nutrient composition data for individual foods (11). The task is substantial, as there are some 10,000 to 15,000 food items in the U.S. food supply and data are being acquired for 60 to 100 nutrients or other food components. Data are collected and entered into the NDB on a continual basis, but the availability of data on some nutrients is limited by the lack of suitable methods. Sources of data include a number of Federal Government laboratories, including USDA's Nutrient Composition Laboratory, university research under Government sponsorship, and food nutrient analyses conducted by industry in support of the nutrition labeling program. The development of the NDB is keyed to the process of revising Agriculture Handbook No. 8, "Composition of Foods: Raw, Processed, Prepared" (12), which is the standard reference table on food composition.

FDA conducts annually its Total Diet Study to estimate average consumption of important components of the diet. This study provides a tracking system for specific indicators of significant changes in the nutritional quality of the national food supply. For example, through this study, FDA has documented the existence of a higher quantity than desirable of iodine in the food supply (13). The quantity of iodine is now being monitored, particularly for some of the food classes noted previously to have high levels, such as milk and cereal grain products (14).

The scientific base for the Total Diet Study has been updated from food consumption data obtained in 1965 to the recently available NHANES II and NFCS data (15). Approximately 200 individual foods are involved, representing about 90 percent of the total foods consumed in the United States. Sampling is done in 30 urban areas in the United States, and analyses for dietary content of seven minerals (iodine, iron, sodium, potassium, copper, magnesium, and zinc) are performed. This is the only extant system for annual chemical analytical measurement of average intakes of pesticides, heavy metals, and environmental contaminants. The study has recently been converted from the measurement of food composites to measurements of nutrients, pesticides, and contaminants in individual foods.

FDA has maintained since 1977 both a surveillance and a compliance program for nutrition labeling. A statistical sample of the 40 percent of processed foods that bear nutrition labels is analyzed for many nutrients on a continuous basis. Annually, approximately 300 foods are analyzed for eight nutrients, involving in excess of 2,000 individual analyses. This surveillance program permits FDA to track the evolution of nutrition labeling in the food supply, assure necessary levels of accuracy of label values, and identify segments of the industry that require encouragement. This activity also permits early identification by FDA of new fortification practices by industry. When combined with consumer studies, reasonable assessments of the value of nutrition labeling are possible.

USDA's Nutrient Composition Laboratory, located in Beltsville, Maryland, provides essential data on the nutrient content of foods consumed in the United States by analyzing the nutrient content of foods with tested, dependable assay techniques; developing either new or improved methods for the analysis of nutrients in foods; developing sound sampling techniques to ensure that representative samples are analyzed; and conducting research on the effect of food-processing procedures and of transportation and marketing methods, as well as home, institution, and restaurant food preparation procedures, on the nutrient composition of foods. Data from these studies are used to update the food composition values in the Nutrient Data Bank.

Dietary knowledge and attitudes

Annually since 1978, FDA has conducted the Public Attitudes Survey based on a national probability sample of food purchasers to measure public attitudes, knowledge, and practices relating to food and nutrition. About one-half of the survey content is concerned with such matters as opinions about nutrition, food quality, and food regulation and is repeated every year for the purpose of tracking changes over time. The other half involves new areas of interest or concern to FDA.

This survey involves detailed interviews in the homes of approximately 1,500 individuals primarily responsible for household food purchases, of whom about 85 percent are women and 15 percent men. Studies of this type permit monitoring of public attitudes and practices about foods and nutrition, as well as identification of public concerns and elements of confusion. It is through this mechanism that the predominance of avoidance practices in food purchases through use of food labels was identified (for example, avoidance of fats and sugars). In addition, assessments are made of the influence of nutrition misinformation and the public's ability to comprehend food label information.

Food supply determinations

Each year since 1909, USDA has calculated the nutrients available for daily per capita consumption from estimates of per capita food availability (retail weight). No deductions are made for waste of food in the distribution system or in the home, use for pet food, or loss of nutrients during the preparation of food. Adjustments are made for nutrients added to the food supply through enrichment and fortification—such as iron, thiamine, riboflavin, and niacin added to flour and cereal products (16).

The estimates of nutrient availability are based on the quantities of 350 foods that "disappear" into the U.S. food marketing system. Hence, the term "disappearance data" is frequently used in referring to these estimates. Levels of food energy (kilocalories) and 15 nutrients (table 2) are calculated using food composition data collected and published by USDA.

Dietary questionnaires used in previous HANES

The five major components of previous Health and Nutrition Examination Surveys were a household questionnaire, a medical history questionnaire, a dietary questionnaire, examination by a physician, and special procedures and tests. The household questionnaire consisted of questions about family relationship; age, sex, and race of family members; housing; occupation, income, and educational level of each family member; and participation in the Food Stamp program and school breakfast and lunch programs. Separate medical history questionnaires were used depending on the age of the sample person, one questionnaire for children 6 months–11 years and another for persons 12–74 years. Both the household and medical history questionnaires were administered in the respondent's home.

When the sample persons arrived at the mobile examination center, they were scheduled through the dietary interview, physician's examination, and special procedures and tests. The procedures and tests for the nutritional assessment included body measurements, urine tests, and blood tests. From blood samples taken in the center, a number of nutrition-related assays were done. These included serum albumin, serum vitamins A and C, serum lipids (cholesterol, triglycerides, and high-density lipoproteins), protoporphyrin, serum iron, total iron-binding capacity, serum zinc, and serum copper. Red cell folates, serum folates, serum ferritin, and serum vitamin B_{12} were determined on blood samples with abnormal complete blood count, hemoglobin, hematocrit, or mean corpuscular volume (MCV) and on a subsample of all other blood samples.

The dietary questionnaires consisted of a 24-hour recall, a food frequency, a supplemental dietary questionnaire, and specific questions on medication and vitamin-mineral supplement usage. All dietary interviews were conducted by trained interviewers who held at least a bachelor's degree in home economics. Copies of the questionnaires used in NHANES II and the Hispanic Health and Nutrition Examination Survey (HHANES) are available (5,6).

During the 24-hour recall, respondents were asked to report all foods and beverages consumed on the previous day. Respondents estimated the size of the portions consumed by referring to food models. In addition to foods and portion sizes, interviewers asked about what time of day the food was eaten and its source. The time of day was coded as one of five ingestion periods—morning, noon, between meals, evening, or total day. The source of the food was coded as home, school, restaurant, or other.

Each food item was coded by the interviewer within 72 hours of the interview. The food codebook developed for NHANES II contained 5-digit food codes for approximately 2,500 food items. Each food item was identified by name, including brand name if appropriate; whether it was raw, dry, canned, or frozen; how it was prepared; and for mixed dishes without food codes, the major ingredients. A food composition data base updated from NHANES I was used to calculate the energy, vitamin, and mineral content of the reported foods. Modifications to the NHANES I data base included new data from USDA's revised Handbook Number 8 (12) and food composition data from food companies on new products and brand name products of unique formulation. The food composition data base for HHANES and future surveys will be the data base in current use by the U.S. Department of Agriculture, which includes 27 dietary components and energy. (See CSFII listing in table 2.)

The questionnaire used to determine the frequency of food consumption has changed with each survey. In NHANES II and HHANES, the food frequency elicited information about the consumption of 18 food groups over the previous 3 months. These groups can be related to the 13 major groups in NHANES I but are progressively more detailed. The frequency was coded as a whole number, never, or unknown. The interval at which the food was usually eaten was coded as never, daily, weekly, or less than weekly. One question was asked about how often the salt shaker was used at the table. Responses to this last question could be assigned to one of three codes—rarely or never, occasionally or seldom, frequently or always.

The supplemental dietary questionnaire contained questions about whether the respondent was on a special diet and, if so, what type and for how long. One question asked about use of medications in the previous week. These were commonly prescribed medications that might interfere with test results or affect interpretation of results. Another question related to problems preventing the respondent from obtaining needed groceries. The final question asked about trouble swallowing, pain, nausea and vomiting following eating, and loss of appetite.

The medication, vitamin, and mineral usage questionnaire has evolved over time. In NHANES II, specific information was requested about the brand name, manufacturer's name, and reason for using vitamin or mineral supplements and medications. Sample persons were left the form and asked to fill it out and bring it with them to the examination centers. Because of low response, the procedures for collecting medication and supplement usage were changed in HHANES. During the household interview, sample persons were asked to bring their medicines and supplements to the interviewer, who recorded the label information.

The quality of the dietary component was controlled at several levels. Before the survey began, the dietary interviewers were trained in interview techniques and in how to code the 24-hour recall. A manual describing the procedures to be followed was issued to each interviewer. Periodically, the forms were reviewed and evaluated, and instructions were issued to the interviewers to promote consistency. Interviewers exchanged coded 24-hour recall forms to check each other's work, and forms were also reviewed by the field staff before forwarding to headquarters. At every location, each interviewer tape recorded two interviews with randomly selected subjects. The recordings were evaluated at headquarters for adherence to procedures. Comparisons were made at headquarters of the mean values and frequency distributions by stand location and by interviewer to detect unusual results by location and systematic errors by interviewers. Foods for which no appropriate food codes existed were forwarded to headquarters for assignment of new code numbers.

Uses and limitations of survey methods

Uses of dietary data

The NHANES dietary data have been put to four types of uses: Relating diet and demographic characteristics, relating diet and health characteristics, determining interactions of diet and nutritional status indicators, and tracking trends in diet and nutrient intakes over time.

In relating diet to demographic characteristics of the population, the major question to be asked is: What are the food consumption patterns and nutrient intakes of subpopulations of the United States by such characteristics as age, race, sex, income, occupation, and education? The NHANES dietary data can answer such questions as: How do nutrient intakes and food consumption patterns of persons differ by level of education? What are the regional differences in consumption of certain food groups?

The NHANES data have been used to relate food consumption patterns and nutrient intakes of subpopulations of the United States to indicators of health status. Specific questions that could be addressed include: How do nutrient intakes compare with the Recommended Dietary Allowances and other dietary guidelines? What dietary patterns and nutrient intakes are associated with differing levels of health or health risk?

Examining interactions among nutrition-related variables, NHANES data could be used to compare among dietary intake, biochemical status, anthropometry, and presence or absence of health conditions. Questions that could be addressed by the data include: What are the relationships among dietary intake, biochemical indicators, and health status for persons who smoke, use vitamin-mineral supplements, or use oral contraceptives? Are those who take vitamins and other dietary supplements the persons who need them? Do the diets of subpopulations with high serum cholesterol levels differ from those with lower levels?

Changes over time in food and nutrient intakes could be tracked and correlations made with health variables. Examples of questions that could be posed to the data include: What changes in obesity, diet, and activity patterns will take place in the next 10 years? Will diet help explain the continuing decline in serum cholesterol values among men and women?

Estimates of individuals' food and nutrient intakes

Of the methods used by Federal surveys, only one would be appropriate for studies of the association between diet and health—estimates of food and nutrient intakes by individuals. National surveys collect information about food and nutrient intakes by individuals for two purposes—to estimate intakes by groups of individuals who share some characteristic (age, sex, race, income, education, and so forth) and to relate estimated intakes by individuals to some condition (such as cholesterol intake and serum cholesterol level).

The national surveys use three methods to collect information from individuals about their food intakes: 24-hour recalls (NHANES, NFCS, and CSFII), food diaries (NFCS), and food frequency questionnaires (NHANES). The 24-hour recall and food diaries attempt to capture actual dietary intakes during specified periods of time; frequency questionnaires attempt to capture usual intakes over longer timespans ranging from a few months to a year. The 24-hour recall and diaries yield quantitative estimates of foods and nutrients ingested. This is not yet possible with food frequencies because a food composition data base has yet to be developed for this purpose. Several reviews (17–19) have appeared about the reliability and validity of these methods. The literature is more extensive for the 24-hour recall method than for diaries or frequencies.

The 24-hour recall and diaries can be used in cross-sectional surveys to determine the foods ingested and to calculate nutrient intakes. Estimates can be made of the mean, median, and distribution of intakes by specific groups described by age, sex, race, income, or other characteristics. The average of repeated recalls and diaries can be used to determine usual food intake by individuals. The minimum number of days of observation required to characterize usual intake varies with the nutrient of interest and by the amount of intraindividual and interindividual variation.

Both the 24-hour recall and the diaries are limited in that they rely on the respondent's ability to remember and describe to an interviewer what he or she ate and drank or to remember to record what was eaten. Probably more frequently than one would like to admit, the diary method becomes a recall when the respondent fails to record what was eaten and relies on memory to complete the diary before it is returned to the interviewer.

The precision of estimates of individuals' usual intakes based on single, 1-day observation is low, largely because of high intraindividual variance (20,21). The major components of variance in the 24-hour recall are sex, day of the week, and interindividual and intraindividual variation. Men tend to eat more than women, so most studies of variability in intakes treat the sexes separately. The day of the week for which the recall was obtained can be a significant source of variance for women, with weekend intakes being higher than weekday intakes. However, the two largest sources of variance are interindividual and intraindividual. Interindividual variations of energy and nutrient intakes are relatively constant, but intraindividual variations differ with the nutrient under consideration. Intraindividual variation is usually larger than interindividual variation and includes methodological errors as well as the true day-to-day variation in intake within respondents. The high intraindividual variability noted in many studies suggests a wide daily variation in the composition of self-selected diets.

The food composition data base used to calculate nutrient levels from reported foods is a contributor to error associated with 24-hour recalls and diaries. The percent of data from analytic sources, that is, actual chemical analysis of food, in USDA's Primary Data Set (which includes nutrient values for all food items needed to create a survey nutrient data base) is shown in table 3. For some dietary components (total dietary fiber, for example) only scanty data exist.

Table 3. Percent of data from analytical sources in Primary Data Set: January 1986

Component	*All food sources*	*Major sources*
Protein	95	97
Fat	95	97
Calcium	90	94
Iron	89	88
Magnesium	74	72
Phosphorus	89	92
Potassium	89	92
Sodium	89	86
Zinc	72	79
Copper	65	70
Vitamin A (international units)	79	94
Vitamin A (retinol equivalents)	60	73
Carotene (retinol equivalents)	54	90
Vitamin E	28	52
Vitamin C	84	95
Vitamin B_1	90	91
Vitamin B_2	90	92
Niacin	89	93
Vitamin B_6	73	70
Folacin	54	68
Vitamin B_{12}	63	69
Cholesterol	80	91
Saturated fatty acids	59	74
Monounsaturated fatty acids	58	78
Polyunsaturated fatty acids	58	72
Total dietary fiber	28	40

SOURCE: Frank Hepburn, U.S. Department of Agriculture. Calculated from the Primary Data Set, which includes nutrient values for all food items needed to create a survey-nutrient data base.

Food frequency questionnaires are useful to classify individuals and groups by food intake characteristics (18). Frequencies can be designed to cover the entire diet (as has been done for NHANES), or they can be targeted to one or more foods or groups of foods of research interest. Information on the frequency of consumption of foods can be combined with estimates of usual portion sizes to yield estimated quantities of foods consumed.

Frequencies suffer from a number of limitations. Because frequencies cover longer periods of time than recalls or diaries, they are more difficult to validate. Usually, validation studies have attempted to compare frequencies with another questionnaire method. As yet, frequencies cannot be used to estimate nutrient intakes. Additionally, like recalls, frequencies rely on the respondent's ability to accurately remember and report his or her usual food intake.

Summary

This paper reviews the strengths, weaknesses, and limits to interpretation of dietary intake methods used in national nutrition surveys, including per capita consumption of food and nutrients and the 24-hour recall. These dietary assessment methods make up but one of five categories of nutritional assessment techniques. The other four are body measurements, hematological and biochemical tests, medical examination for the presence of clinical signs of deficiency or toxicity, and medical history. No one method alone is sufficient for assessing the nutritional status of individuals or groups.

In the hypothesis-generating phases of research on diet and cardiovascular diseases and diet and cancer, correlations of per capita availability of food and nutrients with mortality from these diseases in several countries proved useful. However, the approach is limited to the extent that similar data are available from several other countries and that the association observed is true and not spurious. The 24-hour recall method has proved to provide accurate and reproducible estimates of the mean intakes of population groups, but multiple-day information is necessary for characterizing an individual's usual nutrient intake. Food frequency questionnaires are useful when one is interested in food rather than nutrient consumption of individuals or groups. The 24-hour recall, food diary, and food frequency methods can be used to develop a descriptive epidemiology of food intake within U.S. population groups with specified fitness and activity characteristics or to monitor the prevalence of dietary risk factors within the population. They are limited to the extent that food composition data are available. When selecting a dietary assessment method to be used in a nutrition survey, three points must be kept in mind: Practicality in terms of respondent burden and analysis resources, reliability, and validity.

References

1. Ware JE, Brook RH, Davis AR, Lohr EN. Choosing measures of health status for individuals in general populations. Am J Public Health 71:620–5. 1981.
2. Ostenso G. National Nutrition Monitoring System: A historical perspective. J Am Diet Assoc 84:1181–5. 1984.
3. Miller HW. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–1973. National Center for Health Statistics. Vital Health Stat 1(10a). 1978.
4. National Center for Health Statistics. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–1973. National Center for Health Statistics. Vital Health Stat 1(10b). 1977.
5. McDowell A, Engel A, Massey JT, Maurer K. Plan and operation of the second National Health and Nutrition Examination Survey, 1976–80. National Center for Health Statistics. Vital Health Stat 1(15). 1981.
6. Maurer KR. Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982–84. National Center for Health Statistics. Vital Health Stat 1(19). 1985.
7. Centers for Disease Control. Nutrition surveillance, 1980. Washington: Public Health Service. 1982.
8. Human Nutrition Information Service. Food consumption: Households in the United States, spring 1977. Pub no H–1. Washington: U.S. Department of Agriculture. 1982.
9. Human Nutrition Information Service. Food intakes: Individuals in 48 States, year 1977–78. Pub no I–1. Washington: U.S. Department of Agriculture. 1983.
10. Human Nutrition Information Service. Nationwide Food Consumption Survey continuing survey of food intakes by individuals, women 19–50 years and their children 1–5 years, 1 day, 1985. Pub no 85–1. Washington: U.S. Department of Agriculture. 1985.
11. Hepburn FN. The USDA National Nutrient Data Bank. Am J Clin Nutr 35:1297–1301. 1982.
12. U.S. Department of Agriculture. Composition of foods: Raw, processed, prepared. Agriculture Handbooks 8–1 through 8–12. Washington: U.S. Department of Agriculture. 1976–84.
13. Park YR, Harland BF, Vanderveen JE, et al. Estimation of dietary iodine intake of Americans in recent years. J Am Diet Assoc 79:17–24. 1981.
14. Allegrini M, Pennington JAT, Tanner JT. Total Diet Study: Determination of iodine intake by neutron activation analysis. J Am Diet Assoc 83:18–24. 1983.
15. Pennington JAT. Revision of the Total Diet Study food lists and diets. J Am Diet Assoc 82:166–73. 1983.
16. Welsh SO, Marston RM. Review of trends in food use in the United States, 1909 to 1980. J Am Diet Assoc 81:120–5. 1982.
17. Burk MC, Pao EM. Methodology for large-scale surveys of household and individual diets. Home Economics Research Report No. 40. Washington: U.S. Department of Agriculture. 1976.
18. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 115:492–505. 1982.
19. Bazzarre TL, Myers MP. The collection of food intake data in cancer epidemiology studies. Nutr Cancer 1:22–45. 1979.
20. Beaton GH, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Am J Clin Nutr 32:2546–59. 1979.
21. Todd KS, Hudes M, Calloway DH. Food intake measurement: Problems and approaches. Am J Clin Nutr 37:139–46. 1983.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1992— 3 1 2 - 0 8 2/ 6 0 0 0 1

Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures** – Reports describing the general programs of the National Center for Health Statistics and its offices and divisions and the data collection methods used. They also include definitions and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research** – Studies of new statistical methodology including experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. Studies also include comparison of U.S. methodology with those of other countries.

SERIES 3. **Analytical and Epidemiological Studies** – Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports** – Final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

SERIES 5. **Comparative International Vital and Health Statistics Reports** – Analytical and descriptive reports comparing U.S. vital and health statistics with those of other countries.

SERIES 6. **Cognition and Survey Measurement** – Reports from the National Laboratory for Collaborative Research in Cognition and Survey Measurement using methods of cognitive science to design, evaluate, and test survey instruments.

SERIES 10. **Data From the National Health Interview Survey** – Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, all based on data collected in the continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey and the National Health and Nutrition Examination Survey** – Data from direct examination, testing, and measurement of national samples of the civilian noninstitutionalized population provide the basis for (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

SERIES 12. **Data From the Institutionalized Population Surveys** – Discontinued in 1975. Reports from these surveys are included in Series 13.

SERIES 13. **Data on Health Resources Utilization** – Statistics on the utilization of health manpower and facilities providing long-term care, ambulatory care, hospital care, and family planning services.

SERIES 14. **Data on Health Resources: Manpower and Facilities** – Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

SERIES 15. **Data From Special Surveys** – Statistics on health and health-related topics collected in special surveys that are not a part of the continuing data systems of the National Center for Health Statistics.

SERIES 16. **Compilations of Advance Data From Vital and Health Statistics** – These reports provide early release of data from the National Center for Health Statistics' health and demographic surveys. Many of these releases are followed by detailed reports in the Vital and Health Statistics Series.

SERIES 20. **Data on Mortality** – Various statistics on mortality other than as included in regular annual or monthly reports. Special analyses by cause of death, age, and other demographic variables; geographic and time series analyses; and statistics on characteristics of deaths not available from the vital records based on sample surveys of those records.

SERIES 21. **Data on Natality, Marriage, and Divorce** – Various statistics on natality, marriage, and divorce other than those included in regular annual or monthly reports. Special analyses by demographic variables; geographic and time series analyses; studies of fertility; and statistics on characteristics of births not available from the vital records based on sample surveys of those records.

SERIES 22. **Data From the National Mortality and Natality Surveys** – Discontinued in 1975. Reports from these sample surveys based on vital records are included in Series 20 and 21, respectively.

SERIES 23. **Data From the National Survey of Family Growth** – Statistics on fertility, family formation and dissolution, family planning, and related maternal and infant health topics derived from a periodic survey of a nationwide probability sample of women 15–44 years of age.

SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, Divorce, and Induced Terminations of Pregnancy** – Advance reports of births, deaths, marriages, and divorces are based on final data from the National Vital Statistics System and are published annually as supplements to the Monthly Vital Statistics Report (MVSR). These reports are followed by the publication of detailed data in Vital Statistics of the United States annual volumes. Other reports including induced terminations of pregnancy issued periodically as supplements to the MVSR provide selected findings based on data from the National Vital Statistics System and may be followed by detailed reports in the Vital and Health Statistics Series.

For answers to questions about this report or for a list of titles of reports published in these series, contact:

Scientific and Technical Information Branch
National Center for Health Statistics
Centers for Disease Control
Public Health Service
6525 Belcrest Road, Room 1064
Hyattsville, Md. 20782

301-436-8500

DEPARTMENT OF
HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

DHHS Publication No. (PHS) 92-1464, Series 4, No. 27